Proceedings of
Vidya MCA Departmental Seminar
(VMCADS - 2022)

Vidya Academy of Science & Technology

Thrissur – 680501

Proceedings of

Vidya MCA Departmental Seminar (VMCADS – 2022)

22 – 23 June 2022

Editors

Dr Reji C Joy
Associate Professor of Computer Applications

Sajay K R
Associate Professor of Computer Applications

Siji K B
Assistant Professor of Computer Applications

Department of Computer Applications
Vidya Academy of Science & Technology
(APJ Abdul Kalam Technological University)
Thalakkottukara, Thrissur – 680501, India

PROCEEDINGS OF VIDYA MCA DEPARMTNETAL SEMINAR
(VMCADS -2022)

Distributed by

MCA Department

Vidya Academy of Science & Technology
Thalakkottukara, Thrissur - 680501, Kerala

First Edition

The book was typeset by the editors using the LaTeX document preparation system.

Cover design: Editors

Copyright © 2022 Dr. Saji C. B.

Licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) License. You may not use this file except in compliance with the License. You may obtain a copy of the License at:
`https://creativecommons.org/licenses/by/4.0/`.

Price: Rs. 0.00.

First printing: July 2022

Contents

Preface

1. A Review of Robotics 1
Anjali John, Anjana K U, Jesly Johnson A and Dr Reji C Joy

2. E-Payment System and Its Security in E-Commerce 9
Jineesha Joseph, Jini James, Manju P A and Dijesh P

3. A Perspective Approach on the Cognitive Aspects of Human Computer Interaction 15
Hridhya K B, Sruthy Krishna P U, Swetha M S and Sajay K R

4. An Overview of Cyber Security: Issues, Attacks, Goals and Models 21
Afra P L, Hansa A M, Naina Vincent and Manesh D

5. A Literature Review of Mobile Cloud Computing 27
Anupama P N, Reshma Reghu, Steny Baby and Siji K B

6. Green Computing: The Future of Computing 34
Anjitha M P, Aleena Vincent, Dijina P D and Salkala K S

7. Cloud Computing Architecture 39
Amritha Suresh, Diyana Ummer, Josiya Mary K S and Aparna S Balan

8. A Study on Web Mining Technique : Trends and Applications 50
K R Deepa, Sibin Babu, Sreeharsha V P and Dr Reji C Joy

9. E-commerce Security Issues and Solutions 59
Aysha P A, Dona James, Krishnapriya T S and Dijesh P

10. Health Care, Security Standards and Applications in IoT 64
Arjun K, Lishell Rapheal, Merin V Wilson and Sajay K R

11. Big Data: Analysis and Application 69
Adarsh K Sunny, Frederic Sojan, Rohini T B and Manesh D

12. A Literature Review of Big Data 75
Mariya C Johny, Saranya V Simon, Silpa Sasi and Siji K B

13. Study on Cybersecurity 81
Athira P V, Dhilsha K P, Sandra Thomas and Salkala K S

14. Survey of Machine Learning Applications 87
Devika P H, Shejila K, Meghana K B and Aparna S Balan

15. Image Segmentation and Enhancement: An Overview 95
Adithyalakshmi S Kumar, Jamseena N A, Mrudula M M and Dr Reji C Joy

16. E-Commerce Security 100
Athira K M, Haritha H, Manasa K Mani and Dijesh P

17. A Study on Architecture, Security and Applications of Cloud Computing 107
Ann Mary M B, Dilna K Daison, Mohidh K M and Sajay K R

18. Wireless Sensor Network : Deployment and Security Issues 116
M Shahil Haris, Praseetha Praveen, Suraj Kumar M and Manesh D

19. Cross Platform for Mobile Application Development 122
Blessy Joseph, Janish P S, M R Aswathy and Siji K B

20. Metaverse: The Future of Internet 128
Anjitha N A, Athul P S, Catherine Joshi and Salkala K S

21. A Survey on Need of Green Computing 134
Iyrine George, Maria Benny, Sowmya P N and Aparna S Balan

Name Index 147

Preface

This volume contains the proceedings of the two-day departmental seminar organised by the Department of Computer Applications (MCA) of Vidya Academy of Science & Technology during 04 – 05 July 2022. The seminar was the culmination of a coursework (with course code 20MCA244 Seminar) to be completed by the MCA students of APJ Abdul Kalam Technological University during the Fourth Semester of the MCA programme.

The syllabus of the course 20MCA244 Seminar specifies the objective of the course as follows: "To enable the students to gain knowledge in any of the technically relevant current topics on Computer Science/Information Technology/research, and acquire the confidence in presenting the topic and preparing a report." Again as per the syllabus, as part of the course, "each student is expected to undertake a detailed study on a technically relevant current topic in computer science/information technology under the supervision of a faculty member, by referring articles published in reputed journals/conference proceedings. Each student has to submit a seminar report, based on these papers; the report must not be reproduction of any original paper. The topic has to be presented taking a duration of 15 – 20 minutes. The report and slides for presentation has to be prepared using free typesetting software such as LATEX."

In Vidya Academy of Science & Technology, the supervising teachers helped the students to identify the areas in which the students were to work and the teachers provided the students with some initial learning materials in the form of papers. After the initial reading of these materials, the students were asked to search for additional reading materials themselves. The students were required to study the papers and present a "study report/study paper" in a Departmental Seminar. The reports/papers collected in this volume are the study papers prepared by the students and presented in the Departmental Seminar. As part of the learning process, the students were also required to present the paper in the IEEE conference paper format. To facilitate this, the students were given a basic introduction to the LATEX software and the IEEEtran document style.

In addition to gaining knowledge in any of the technically relevant current topics on Computer Science/Information Technology/Research, the course also aimed at giving the students a hands on experience in preparing a conference/seminar paper. The expected learning outcomes of the course included acquiring a clear knowledge about the following aspects of preparing and presenting a high quality research paper:

- The structure of a research paper
- The process of literature survey
- The accurate preparation of bibliography and their citations in the paper
- The IEEE format for the preparation of conference/journal papers
- The concepts of "Abstracts", "Keywords", and the like

- The methodology of presenting a multi-author paper in a seminar/conference.

The articles compiled in this Proceedings are not even moderately edited. The editors have only ensured that the basic learning outcomes outlined above have been met. However, the editors have tried to ensure that the titles of chapters, sections, etc., the abstract, figure and table captions, and the like are as per IEEE guidelines. The references have not been checked for accuracy and completion. The papers have not been edited for grammar, punctuation, spelling or style.1

The present work is only a record of the activities of the course referred to above and it is prepared only for private circulation. To the best of our understanding the authors of the papers have given proper attribution to ideas and material presented in the papers. If there are no attributions or improper attributions, it was unintentional. Hence the contents have not been subjected to plagiarism tests. It is believed that the teachers as well the students have greatly enjoyed doing the seminar course. There are still much scope for improvement. It is our hope that the future batches of students will have a stronger and wider learning experience from similar seminar courses.

Editoral Team
Proceedings of
VMCADS – 2022

Proceedings of
Vidya MCA Departmental Seminar (VMCADS – 2022)

A Review of Robotics

Anjali John, Anjana K U,
and Jesly Johnson A
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dr Reji C Joy
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: reji.c.j@vidyaacademy.ac.in)

***Abstract*—The purpose of this paper is to present an analysis on robotics. Robotics is a branch of engineering that involves the conception, design, manufacture and operation of robots. The objective of the robotics field is to create intelligent machines that can assist humans in a variety of ways. In this paper we look into the usefulness of Artificial Intelligence in robotics. Artificial intelligence is a field of study and the resulting innovations and developments that have culminated in computers, machines, and other artifacts having human-like intelligence characterized by cognitive abilities, learning, adaptability, and decision-making capabilities. The study ascertained that AI has extensively been adopted and used in education, particularly by education institutions, in different forms. The Cloud infrastructure and its extensive set of Internet-accessible resources has potential to provide significant benefits to robots and automation systems is another aspect of this paper. Human Robot Interaction (HRI) is an application of robotics that has recently received considerable attention in the academic community, in labs, in technology companies and through the media.**

***Index Terms*—Robotics, Artificial Intelligence, Human Robotic Interaction(HRI), Cloud Robotics.**

I. Introduction

Today, it is increasingly common for people to use or come into contact with robots in various situations at home, retail stores, hotels and hospitals. Robots are classified into several types based on their functionality (service and utility robots or those designed to communicate with humans) and appearance (humanoid robots or mechanical robots). The type of robot, to which each country attaches particular importance in the advance of robotics, reflects the sense of values and preferences of its population. Where the U.S. has the highest level of robot utilization at home and in retail stores with its people being the most enthusiastic about the future use of robots. Germany shows a strong tendency to consider robots for industrial purposes, and its people feel strong resistance to the presence of robots in their households. Robotics is anything but a static field with a continuous stream of advancements adding to both the complexity and possibility behind each new development. A new subfield–"cloud robotics"—is emerging as a hot topic in research. As one might imagine, instead of relying on "in-house" resources, robots can potentially leverage the cloud to deliver instant information and to handle computationally intensive tasks that would otherwise use a great deal of a robot's on-board system. For us humans, with our offline, non-upgradable meat brains, the possibility of acquiring new skills by connecting our heads to a computer network is still science fiction. Not so for robots. Cloud computing, one of the most promising technologies, when coupled with robotics can make huge changes in their abilities. Although robot technology was primarily developed in the mid and late 20th century, it is important to note that the notion of robot-like behaviour and its implications for humans have been around for centuries in religion, mythology, philosophy, and fiction. The word "robot" originates from the Czechoslovakian word 'robota' which means 'work'. "Robot" appears to have first been used in Karel Chapek's 1920's play Rossum's Universal Robots, though this was by no means the earliest example of a human-like machine. Early robot implementations were remotely operated devices with no or minimal autonomy. In 1898, Nicola Tesla demonstrated a radio-controlled boat, which he described as incorporating "a borrowed mind." In fact, Tesla controlled the boat remotely. Complementing the advances in robot mechanics, research in artificial intelligence has attempted to develop fully autonomous robots. The most commonly cited example of an early autonomous robot was Shakey, which was capable of navigating through a block world under carefully controlled lighting conditions at the glacially slow speed of approximately 2 meters per hour. A breakthrough in autonomous robot technology occurred in the

Fig. 1. Google's autonomous car

mid-1980s with work in behaviour-based robotics. Another important breakthrough for autonomy as it applies to Human Robot Interaction (HRI) is the emergence of hybrid architectures; these architectures simultaneously allow sophisticated reactive behaviours that provide fundamental robot capabilities along with the high-level cognitive reasoning required for complex and enduring interactions with humans.

II. Cloud Robotics and Automation

The Cloud has potential to enable a new generation of robots and automation systems to use wireless networking, Big Data, Cloud Computing, statistical machine learning, open-source, and other shared resources to improve performance in a wide variety of applications including assembly, inspection, driving, warehouse logistics, care giving, package delivery, housekeeping, and surgery. Cloud Robot and Automation systems can be broadly defined as follows: Any robot or automation system that relies on either data or code from a network to support its operation, i.e., where not all sensing, computation, and memory is integrated into a single stand alone system. Consider Google's autonomous car. It uses the network to index maps, images, and data on prior driving trajectories, weather, and traffic to determine spatial localization and make decisions. Data from each car is shared via the network for statistical optimization and machine learning performed by grid computing in the Cloud.

Another example is Kiva Systems approach to warehouse automation and logistics using large numbers of mobile platforms to move pallets using a local network to coordinate platforms and share updates on floor conditions. Google's James Kuffner coined the term "Cloud Robotics" in 2010. Cloud Robot and Automation systems can be broadly defined as any robot or automation system that relies on data or code from a network to support its operation, i.e., where not all sensing, computation, and memory is integrated into a single stand alone system. There are at least four potential advantages to using the Cloud:

1) Big Data "Big Data" describes "data that exceeds the processing capacity of conventional database systems" including images, video, maps, real-time network and financial transactions, and vast networks of sensors.
2) Cloud Computing Cloud Computing has potential to speed up many computationally intensive robotics and automation systems applications such as robot navigation by performing SLAM in the Cloud and can access to parallel grid computing on demand for statistical analysis, learning, and motion planning.
3) Collective Learning The Cloud facilitates sharing of data for robot learning by collecting data from many instances of physical trials and environments. For example, Lightning path planning framework which proposes a framework for collective robot learning by indexing trajectories from many robots over many tasks and using cloud computing for parallel planning and trajectory adjustment.
4) Human Computation : crowdsourcing and call centers Human Computation use of crowdsourcing to tap human skills for analyzing images and video, classification, learning, and error recovery.

In May 2001, the IEEE Robotics and Automation Society established the Technical Committee on Networked Robots which organized a number of workshops. Chapters of the Springer Handbook on Robotics were focused on Networked Tele-robots and Networked Robots respectively. In 2009, the RoboEarth project was announced. It envisioned "a World Wide Web for robots: a giant network and database repository where robots can share information and learn from each other about their behaviour and environment" as illustrated in Under a major European Union grant, the RoboEarth research team developed a series of system architectures for service robotics developing Cloud networking and computing resources to generate 3D models of environments, speech recognition, and face recognition. Cloud Robotics and Automation is related to several other new initiatives, such as "Internet of Things" a term is introduced in 2010, "Industry 4.0," introduced in Germany in 2011 and in 2012, General Electric also introduce the term "Industrial Internet".

A. Five Traits of Performance Enhancement using Cloud Robotics

In improvements in communication technology and the increasing availability of network, new approaches where robot and automation processing are performed remotely with access to large scale datasets, support a range of functions. Cloud Robotics supplements performance enhancement of robotics and autonomous systems by providing a global infrastructure in innovative ways. This paper summarizes recent research into five traits of Cloud Robotics for performance enhancement in robotics and autonomous systems: This paper summarizes recent research into five traits of Cloud Robotics for performance enhancement in robotics and autonomous systems:

1) Remote Brain
2) Big Data and Shared Knowledge-base
3) Collective Learning
4) Intelligence and Behaviour
5) Cloud architectures

1) Remote Brain (Offloading Computation) :
Robots are typically capable of processing information using multiple onboard processors. Certain types of robots such as mobile robots or Unmanned Aerial Vehicles (UAV) may have limited capacities of onboard processing due to their stringent requirements of dimensional size and limited payload. Such robots can harness remote processing by offloading heavy computation tasks to external computers. Cloud computing enables remote computation by enabling massively parallel computation on demand. Public domain cloud service providers such as Amazon Web Services Elastic Compute Cloud, known as EC2, Google Compute Engine, and Microsoft Azure allow computing resources to be rented

for various computing tasks. Researchers are increasingly utilizing these services for scientific and technical high-performance computing (HPC) tasks. The concept of HPC as a Service, aiming to transform traditional HPC resources into a more convenient and accessible service. Various recently published works have utilized Cloud computing as a platform for offloading and processing compute tasks remotely. Agostinho developed a cloud platform for robotics workflows using the Platform as Service approach. They concluded that the proposed platform provides strong advantages in performance without compromising security.

2) Big Data and Shared Knowledge-base
The sheer scale of data that cannot be managed, transformed, moved or analyzed efficiently without using high performance computing resources, according to F. Alexander, it is termed as Big Data. Examples of Big Data typically involve data that may not be structured using a relational database, including videos, images, maps, graphs, financial transactions, and real time network traffic. Autonomous robots are typically equipped with multiple sensors, cameras etc. Data collected from these sensors is normally used for taking decisions on localization, movement, path determination etc for the robot. Sharing this data among multiple robots increases the volume to an extent that managing it would overwhelm limited capacities onboard a robot. To this end, knowledge bases for all collected may have to be created for processing large scale information in further research. Various research projects have yielded a large collection of datasets and knowledge bases that are available online. The Columbia Grasp dataset 30 is a large dataset of pre-computed grasps on thousands of 3D models, that have been used in many research studies. Application of robotics in medical and healthcare industry is also gaining enormous attention. One such example is Yokoo's work where physical therapy robots collect patient information and log it into a medical cloud. This information serves as a knowledge-base for robots and medical practitioners in learning and improving the medical processes.
3) Collective Learning
Cloud computing provides an excellent platform for networked robots to share data, process information and improve perception by various machine learning methods. A simple example of collective learning can be inspired from a disaster recovery scenario where various mobile robots are deployed to build a map of the location. Many robots communicating through a central cloud service, can post localization information that can be shared by other robots certainly enhancing map discovery. Furthermore, robots can also post control policies, sensor information of physical traits of an environment, trajectories, motion coordination, tracking data and updated localization data for collective learning. Ubiquitous Networked Robot Platform (UNR-PF) as a framework for distributed task coordination and control. The UNR-PF deals with distributed task execution and supervision on multiple robots and sensor zed devices at different locations.
4) Intelligence and Behaviour
Data mining into data-sets and history of trajectories for data obtained by various sensors on the robot can lead to improvement in intelligence and behaviour of autonomous systems. Recently A. Lim highlighted the issue of robot behaviour in human-robot interaction. Three signs including spoken language synthesis, gesture and gaze play a key role in determining metrics and parameters for human-robot interaction. In a cloud robotics context, data available in various shared datasets and knowledge-bases can be mined for improvement in robot intelligence and human-robot interaction and behaviour. Researchers is design a Cooperative human robot interaction system, a cognitive system with focus on gaze and speech. The robot captures speech and gesture from humans, processes this information and produces accurate robot interaction. Researchers tested their system on 12 subjects in a cooperative task and reported quality of cooperation in terms of proper turn taking, collisions and cognitive errors.
5) Cloud Architectures
Cloud provides scalable resources for sharing information between humans, systems and robots. Recently many works have provided middle-wares, frameworks and platforms for Cloud robotics. Nakagawa developed a distributed service framework that realizes the coordination of various devices, robots, and service functions based on RSNP (Robot Service Network Protocol), a protocol specification for robot services. Kato propose the Research Cloud (RSi-Cloud), which enables integration of robot services with Internet services and mash-up of robot services. A Researchers is address Next Generation Cloud Computing and propose Architecture for cloud robot networks that allow sharing of information among various mobile robots. Riazuelo demonstrate a cloud framework for cooperative tracking and mapping (C2TAM); computer intensive bundle adjustment for simultaneous localization and mapping (SLAM) performed in the Cloud. Koubaa proposed RoboWeb, a SOAP-based service-oriented architecture that virtualizes robotic hardware and software resources and exposes them as services through the Web, contributing to the evolving concept of cloud robotics. RoboWeb consists in the integration of different Web services technologies with the (ROS) middleware to allow for different levels of abstraction (multi-layer architecture), ensuring more modularity and flexibility of the deployment.

III. AI IN ROBOTICS

Artificial Intelligence and Robotics have a common root and a (relatively) long history of interaction and scientific discussion. The birth of Artificial Intelligence and Robotics takes place in the same period (‘50) and initially there was

no clear distinction between the two disciplines. The reason is that the notion of ”intelligent machine” naturally leads to robots and Robotics. One might argue that not every machine is a robot, and certainly Artificial Intelligence is concerned also with virtual agents (i.e. agents that are not embodied in a physical machine). On the other hand, many of the technical problems and solutions that are needed in order to design robots are not dealt with by Artificial Intelligence research. According to surveys conducted by the Consumer Electronics Association (CEA) in 2013 and 2014, many American consumers show a certain degree of understanding for task-specific usage of robots. However, the usage in this case is limited to relatively simple tasks and consumers show much less interest in tasks involving interaction and communication. Specifically, more than 50 percent of consumers are interested in having a robot help with household chores, gardening and lawn care, and home security. One example of household robots, which has gained popularity in Japan, is **“Roomba,”** a series of floor-cleaning robots manufactured and distributed by iRobot Corporation. Another is the robotic lawn mower “Rob mow” being sold in the U.S. and the U.K. Nonetheless, when it comes to leisure, time spent with friends or caring for pets and children, American consumers are noticeably less interested in being involved with robots.

1) Shopping assistant robots
   Lowe’s, a U.S. chain of retail home improvement and appliance stores, has partnered with Fellow Robots, a Silicon Valley start-up, to install a retail service robot named “OSHbot” at its hardware supply stores in San Jose, California. In this trial, when a customer tells OSHbot what he or she is looking for, OSHbot provides information about the product and its inventory, and even guides the customer to the item’s location within the store.

Fig. 2. OSHbot

2) Hotel room service robots
   Similar trials are also taking place in hotels. Savioke, which is also a Silicon Valley start-up, developed a room service robot named “Relay,” which has the ability to autonomously deliver new towels, toothbrushes and tubes of toothpaste to guest rooms on behalf of hotel staff when guests call for room service.

Fig. 3. Relay

3) Telemedicine robots
   Since 2012, iRobot, in partnership with InTouch Health, a U.S. telehealth network provider, has been operating its telemedicine robot “RP-Vita” equipped with a camera, microphone and display screen. The robot enables a doctor to move anywhere in the hospital and consult with patients remotely.

Fig. 4. RP-Vita

Both in Japan and in the U.S., the majority of robots being sold in the market are ones that perform specific functions such as **“Roomba,”** the floor cleaning robot previously mentioned. On the other hand, robots that communicate with humans are called “social robots,”whose capabilities continue to improve as AI-related breakthroughs are seen in the area of deep learning. Furthermore, due to increased media coverage, social robots have quickly become well recognized among consumers. For example, in Japan, SoftBank Robotics’ humanoid robot **“Pepper”** has already been introduced by Nestlé,

Mizuho Bank and SoftBank mobile phone stores as well as by Benesse, an education and daycare service provider.At the same time, Aldebaran Robotics' humanoid and programmable robot **"Nao"** is being in- stalled at Bank of Tokyo-Mitsubishi UFJ branches, while Fujisoft's communication robot **"Palro"** is being adopted by nursing facilities to help staff members. The robots introduced thus far, including social robots, can be classified into several types based on their functionality and appearance. By using this classification framework,there are three main groups of robots. The first group includes utility robots with a mechanical appearance such as "Roomba" and Amazon's voice-activated smart speaker robot, "Echo." The second group includes "Jibo" and "Buddy" that are not humanoid robots but are categorized as social robots because of their ability to communicate with humans. American and European manufacturers have taken the lead in the development of robots in this group. The third group consists of robots with human-like appearance such as "Pepper" that also communicate with humans. Many of the robots that belong to this category are being manufactured by Japanese companies. Although Aldebaran Robotics, which develops and markets "Nao," is originally a French company,the fact that it is a subsidiary of SoftBank in Japan means the third group as a whole consists almost entirely of Japanese products.

Fig. 5. Roomba

IV. AI IN EDUCATION

Artificial intelligence is a field of study and the resulting innovations and developments that have culminated in computers, machines, and other artifacts having human-like intelligence characterized by cognitive abilities, learning, adaptability, and decision-making capabilities. The study ascertained that AI has extensively been adopted and used in education, particularly by education institutions, in different forms. AI initially took the form of computer and computer related technologies, transitioning to web-based and online intelligent education systems, and ultimately with the use of embedded computer systems, together with other technologies, the use of humanoid robots and web based chatbots to perform instructors' duties and functions independently or with instructors. Using these platforms, instructors have been able to perform different administrative functions, such as reviewing and grading students' assignments more effectively and efficiently, and achieve higher quality in their teaching activities. Prior to the introduction of computers and other related technologies, instructors and students, engaged in instructions and learning mechanically, or through the pure application of natural human effort. With the introduction of microcomputers, and by extension, personal computers in the 1970s, which according to Flamm, provided more computing power and marked an important transition to electronic computers for them as market Later developments in computers and computer-related technologies, including networking, the internet, the world wide web, and increased processing, computing, and other capabilities, including different programs and software packages that are task oriented, have seen the increased application of computers in different ways in the education sector. More specifically, in different departments in education institutions.

- Education Administration
  AI application in education, in its various forms and serving different functions, has had a major impact on the performance of administrative and management functions in education. It has enabled instructors or teachers to perform their administrative functions, such s as grading and providing feedback to students more effectively. AI has made the performance of administrative tasks easier and improve teacher or instructor efficiency and effectiveness in providing instructions and guidance to students.
- Instruction
  Another aspect of education that was the focus of this analysis is the use of AI in instructions or by instructors. An analysis of different articles showed rapid uptake and use of AI, in different forms for instructional purposes or as a pedagogical tool by instructors. The use of AI for instructional purposes or as a pedagogical tool has had a major impact on this aspect of education. It has improved effectiveness, efficiency, and quality of the work done by instructors as adduced from the different publications reviewed and analyzed. Efficiency and quality, within this context, is measured by the delivery of the relevant content in line with the curriculum and in line with the learner specific needs and capabilities, while effectiveness is assessed by the implied uptake and retention or the achievement of learning by the students or the learners.
- Learning
  Another area of education within the scope of this study, which has majorly been impacted by the rollout and the use of AI, is students' learning experiences. AI has enabled the development and use of intelligent learning systems and adaptive content customized for each students learning needs and capabilities, such as intelligent virtual reality and use of the same in simulation teaching and learning, which has been shown to have a positive impact on learning. Other benefits of AI to learning, and impact on learning are highlighted in other studies.

V. HUMAN–ROBOT INTERACTION (HRI)

HRI-Human-Robot Interaction has the attention seeker in the academic community in labs and other various technologic

companies thorough the media during these era. Therefore HRI surveys serve as an awareness to the people who are outside and make more update to those who are within the field. One of the major aim of the review is to present an ethnic treatment to the problems related to HRI, to identify key themes and also discuss about what are the main challenges that would help us to shape the field in the near future.

A HRI problems means is to understand and shape the interactions between one or more humans and one or more robots. So therefore there are some essential components for HRI. Like evaluating the capabilities of robots training, designing new and new technologies. Here we describe the HRI story from different perspectives. The survey includes papers by scholars from various universities, government efforts, industry labs and also fair reports from various countries that contribute to HRI.

To the core Human-Robot Interaction is the field that cover under the topics like understanding designing and evaluating robotic system by humans. In other words we can say that, communication between the human and robot was closely related to they know each other or not. Mainly there interaction can be separated into two general categories. Remote interaction and proximate interaction and these interactions are physical interactions. Social interactions will include, cognitive aspects and emotive interactions. Here the humans and robots will act as companions.

Through this review we also introduce the modern HRI. Modern HRI is a field with strong social implications. So we will closely identify what are challenges or problems faced by modern HRI. Including now humans and machines interact each other and how much is the collective. Before that we want to know how and when the robotics established. So we will discuss about the early history of robotics. The establishment of robots was a remarkable invent in the mid and late 20's century. Its features of humans came from centuries in mythology, philosophy, religion and fiction. The word 'Robot' originates from a Czechoslovakian word 'Robota' which means work. Karel Chapek was the first person who use robots in his play 'Rossum's Universal Robots'. It was the first example of human like machines.

Fig. 6. Rossum's Universal Robots

Later Leonardo Da Vinci also sketched robot like human or mechanical man in 1495. Then invention of humanoid robots also with the help of many ancient countries like China, Eygpt, Greece. The creation of humanoid exactly work like human being and movement also. Later these were widely seen through the literature and other fictions by various famous writers like W.B.Yeats one among them was the most notable work is the Azimou's works ,the Asimou's laws of robotics was considered as the first designer guidelines of HRI. However the invention in robot machines helped to develop fully autonomous robots. But there are some drawback to it. Mostly these robot are behavior based robotics. But these robot collapse with the monolithic sense plan act loop instead of that they use distributed sense response loops.

The others unavoidable contributions was from international space agencies and NASA. NASA developed robust robot platforms and communications is a vital role in developing HRI. The space researchers made several research and planets according to this to exploring the planets and moons. Because most of their failures occur due to the failure of machines or software problems. So they are very much concerned about robots and research to develop more and more efficient machines or robots. Form these early works the experts focuses on two major paradigms for human-robot interaction they are teleoperation and supervisory control. In teleoperation we can easily control robots by remotes. But in supervisory control there should be human supervise the autonomous system and intervenes.

Various research where there according to HRI, in various fields like programing, refining, testing and so on. The multidisciplinary field is also considered as a part of HRI. ACM International conference on human robot interaction created special attention to multidisciplinary aspect of HRI research in 2006. Competitions made more influence in the emergence of HRI during these era. There two greatest competitions, the one is the AAAI Robotics competition and exhibition and the other one is the Robocup search and rescue competitions. They main attraction of the competition was that the robots will serve snacks to the members of the conference during the conference reception. How the competition occurs was they set a team of robots in front of the door of the venue. Then the robots wants to interact with the people around there and wants to find the way for the registration desk. Then robots itself register for the conference and they wants to find its way to the exact place in the time limit; where the presentation takes place. Like that there are several task assigned for them. The task will be of different modes. It may include effective social interactions, natural interactions like gestures and speech and emotive computing. Space robotics had a vital role in HRI. Because there are some challenges rise under some conditions like while operating remote robot or in proximity.

HRI can roughly divided into four factors of applications .The first one is the Human supervisory control of robots in performance of routine tasks .This section include manufacturing, accessing, delivering. These machines are known as Telerobots, which is capable of carrying out a limited series of actions automatically, based on a computer program. Through these program they can sense their environment, identifying

their positions and also can communicate with people back according to their updates .The second one is the remote control of space. These machines which work under the control of remote are called teleopetators. The third one was the automated vehicles in which a human is a passengers. The final one is the Human robot social interaction. These include teaching, entertainment, comforting, assisting handicapped persons, elderly and children. These four sections explains the functions of robots. There are a huge variety of robots. Therefore there are so many challenges to this area. The human supervisory control of robots are an essential factor in HRI. These human control functions includes teaching, planning, monitoring of automatic control. Human robot social interaction is the other factor. We use programs to process of the face and voice of another person. Kismet makes it possible to show the gestures in return. According to some research these machines were effective is interacting with young children. So we can understand that such devices are good in healthy imagination. Using these ideas there are several toys, like Barbie doll which can speak and tries to recognize voices. There are so many items which are available in our Markets nowadays. One of the important factor of human factors tool kit is task analysis. There will be a difference in humans and automation. This will find the difference between man and machine. Check whether man do or machine do the best. We can teach a robots by giving geometric instruction to how to move, how to avoid, when to move, what to avoid. Teaching robots means by symbols is much easier than teaching analytical language. Robots had a vital role in education. They can easily scan a book and lectures but it is difficult to work with children and adults. They can't able to pop up with them. Nowadays we can see robots everywhere in movies like star wars, wall -E .Human robot interaction is a growing field of research and application. Robot are there in every field like education, research so on. So that human involvement decreases very slowly computers and robot are evolving at a very immediate.

VI. Pros and cons

Robotics possess enormous potential to solve social issues such as labour shortages and improve productivity in a broad range of industries from manufacturing to medical service and nursing care to agriculture, construction and infrastructure maintenance. While it remains some way off to see self driving cars running freely on public roads, some autonomous driving technologies are already at the stage of practical use. In the not too distant future, we will be relieved from some routine driving tasks and the constraints associated with driving. Looking at the other side of the coin, such technologies could end up taking over most jobs from humans. Jobs that are most susceptible to computerization in Japan include train engineers, accounting clerks, tax preparers, mail clerks,taxi drivers and receptionists. On a positive note, importance should also be placed on the provision of opportunities to people to acquire technological expertise essential for the utilization of robots and AI solutions. It would be essential to provide consumers with appropriate knowledge and information so that they can recognize what is necessary and what is not.The utilization of robots and AI is no longer a quaint science fiction concept. Rather, such technologies have great potential to change our ways of living and eventually, our society as a whole. On the other hand, the survey also disclosed their negative attitudes in terms of factors such as the appearance and tasks performed by robots.The time has come for us to seriously think about how we should respond to the technology surrounding robots and AI considering the consumers' sense of values in each country.

VII. Conclusion

The objective or the purpose of this study was to present an analysis on robotics. A qualitative research study, leveraging literature review as a research design and method was used. Journal articles, professional publications, and professional conference reports were identified and used in an analysis that facilitated the realization of the study purpose. Here we conclude that using the Cloud for robotics and automation systems introduces many new challenges. The connectivity inherent in the Cloud raises a range of privacy and security concerns. Cloud Robotics and Automation also introduces the potential of robots and systems to be attacked remotely: a hacker could take over a robot and use it to disrupt functionality or cause damage. On the technical front, new algorithms and methods are needed to cope with time-varying network latency and Quality of Service. Faster data connections, both wired internet connections and wireless standards such as LTE are reducing latency, but algorithms must be designed to degrade gracefully. Moving robotics and automation algorithms into the Cloud requires frameworks that facilitate this transition. The Cloud provides three possible levels at which a framework could be implemented. The lowest level is Infrastructure as a Service (IaaS), The second level, Platform as a Service (PaaS), and at last Software as a Service (SaaS).

AI in education initially took the form of computers and computer related systems, and later, the form of web-based and online education platform. Embedded systems have made it possible to use robots, in the form of cobots or humanoid robots as teacher colleagues or independent instructors, as well as chatbots to perform teacher or instructor-like functions. The use of these platforms and tools have enabled or improved teacher effectiveness and efficiency, resulting in richer or improved instructional quality. Similarly, AI has provided students with improved learning experiences because AI has enabled the customization and personalization of learning materials to the needs and capabilities of students.

Human–robot interaction is a growing field of research and application.The field includes many challenging problems and has the potential to produce solutions with positive social impact. Its interdisciplinary Nature requires that researchers in the field understand their research within a broader context.Research in the areas relating to lifestyle, Fears, and human values is probably the most important challenge for HRI. If a job can be more efficiently done by a robot, should that job always be automated? How can robots be made To

work as active, intelligent task helpers in industrial tasks or in caring for the sick or elderly?In national security contexts, should robots ever be authorized to "use their own judgment" in deciding whether to kill someone or destroy property, or must humans always be in the decision loop? Many writers are posing these big Questions, but research is lacking. Can human factors methods and insights be extended to help answer such questions, or do we have nothing to Contribute?

REFERENCES

[1] Ben Kehoe,Sachin Patil,Pieter Abbeel,Ken Goldberg, "A Survey of Research on Cloud Robotics and Automation",IEEE Transactions on Automation Science and Engineering,Volume: 12, Issue: 2,Page(s): 398 - 409, April 2015

[2] Hiroyuki NITTO,Daisuke TANIYAMA,Hitomi INAGAKI,"Social Acceptance and Impact of Robots and Artificial Intelligence",Findings of Survey in Japan, the U.S. and Germany,February 2017

[3] Michael A. Goodrich,Alan C. Schultz,"Human–Robot Interaction: A Survey",Foundations and Trends in Human–Computer Interaction Vol. 1, No. 3 (2007) 203–275,2007

[4] K.Flamm,Creating the Computer: Government, Industry, and High Technology. Washington, DC, USA: Brookings Institution Press, 1988.

[5] M.Campbell-Kelly, Computer, Student Economy Edition: A History of the Information Machine. Evanston, IL, USA: Routledge, 2018.

[6] M. M. L. Cairns "Computers in education: The impact on schools and classrooms," in Life Schools Classrooms. Singapore: Springer, 2017, pp. 603– 617.

[7] B. Coppin, Artificial Intelligence Illuminated. Boston, MA, USA: Jones and Bartlett, 2004.

[8] B. Whitby, Artificial Intelligence: A Beginner's Guide. Oxford, U.K.: Oneworld, 2008.

[9] ARTIFICIAL INTELLIGENCE AND ROBOTICSANTONIO CHELLADipartimento di Ingegneria Infor matica - Università degli Studi di Paler m oLUCA IOCCHIDipartimento di Infor matica e Sistemistica - Università degli Studi di Roma "La Sapienza" IRENE MACALUSODipartimento di Ingegneria Infor matica - Università degli Studi di Paler m oDANIELE NARDIDipartimento di Infor matica e Sistemistica - Università degli Studi di Roma "La Sapienza"

[10] "For Big-Data Scientists, Janitor Work Is Key Hurdle to Insights." [Online]. Available: http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html

[11] Erico Guizzo,"Robots with their heads in the clouds",IEEE Spectrum, Volume: 48, Issue: 3,Page(s): 16 - 18, March 2011

[12] Antonio chella,Artificial Intelligence and Robotics.[Online].Available: https://www.researchgate.net/publication/220672741 Artificial Intelligence and Robotics

[13] Katie Terrell Hanna,Robotics[Online].Available: https://www.techtarget.com/whatis/definition/ robotics text=Robotics

[14] "IEEE Networked Robots Technical Committee." [Online]. Available: http://www-users.cs.umn.edu/isler/tc/

[15] "What is RoboEarth?" [Online]. Available: http://www.roboearth.org/

[16] G. Hu, W. Tay, and Y. Wen, "Cloud Robotics: Architecture, Challenges and Applications," IEEE Network, vol. 26, no. 3, pp. 21–28, 2012.

[17] "Google Goggles." [Online]. Available: http://www.google.com/mobile/goggles/

[18] S. Gammeter, A. Gassmann, L. Bossard, T. Quack, and L. Van Gool, "Server-side Object Recognition and Client-side Object Tracking for Mobile Augmented Reality," in IEEE Conference on Computer Vision and Pattern Recognition, June 2010, pp. 1–8.

[19] Thomas B. Sheridan,"Human–Robot Interaction: Status and Challenges",HUMAN FACTORS Vol. 58, No. 4, June 2016, pp. 525–532,2016

[20] Sheridan, T. B. Telerobotics, "automation and human supervisory control." Cambridge: MIT Press,1992

E-Payment System and Its Security in E-Commerce

Jineesha Joseph, Jini James and Manju P A
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dijesh P
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: dijesh.p@vidyaacademy.ac.in)

***Abstract*—E-commerce provides the capability of buying and selling products, information and services on the Internet and other online environments. The ecommerce sector has boomed with the ease of purchasing and selling products or services, availability and connectivity of latest technologies over the Internet. As the trust level towards ecommerce platforms are increasing; the consumers are becoming progressively comfortable in using different applications like carts, catalogs, comparison, making online transactions. Although many consumers are getting tuned towards usage of various electronic devices and E-payment platforms by installing and using numerous applications in their smart devices but there are still many consumers who prefer Cash on Delivery. So, it is important for the online businesses specifically to understand the payment usage of ever-changing consumers for better sustainable growth. In an e-commerce environment, payments take the form of money exchange in an electronic form, and are therefore called Electronic Payment. E-Payment system an integral part of electronic commerce. Electronic payment is apparently the key for the online transactions, so its security issue becomes the center of focus in the e-commerce development. Privacy and Security are the two major factors that affect costumers trust in electronic transaction. On this research is a try to make e-commerce system greater secured and save you the unauthentic operations. This article intends to have an in-depth discussion of E-payment methods, Digital wallets and various security measures and protocols which are used for the security of online transaction.**

***Index Terms*—E-commerce, E-payment, electronic transactions, Security, Digital Wallet**

I. Introduction

In today's digital era the usage of the internet has increased drastically. All counts of age are consuming and producing loads of data each second. Many organizations are analysing and altering based on the study that they been able to capture from the consumer analysis. With the ease of smartphones and virtual accessibility of products, services and payments online have boosted the way people are shopping and making payments online. We all have witnessed that the traditional payment modes have been replaced by various types of e-payments that are quick and efficient. In e-payment process both buyer and seller use digital modes to send or receive money, it is an automatic process where seller and the buyer can avoid visiting their bank. It eliminates the physical cash that is risky to handle at times. Today consumers can make payment through electronic modes by using cards and other platforms that are made available through all types of smart devices. The acceptance of particular payments by sellers and businesses has an important influence on the purchases made by their clients, as the availability of the various modes help the consumer to choose that is suitable or available to them as well.

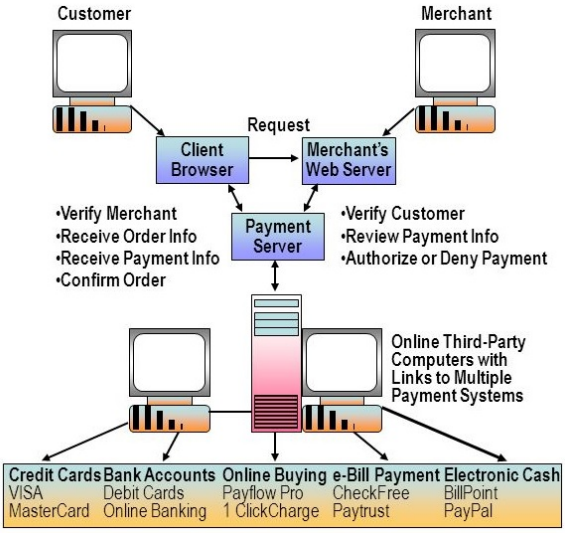


Fig. 1. Payment process

In real world we have two distinct types of payment systems:

- Internet–Based payment system :
  There are four models of Internet-Based payment system:
  - e-Cash
  - Credit Card
  - Debit Card
  - Smart Card
- Electronic Transaction-Based payment system:
  - Secure Electronic Transaction
  - Cyber Cash
  - Net Bill
  - First Virtual Holdings

E-Payment system is secure. There should be no threat to the user credit card number, smart card or other personal

detail, payment can be carried out without involvement of third party, it makes E-payment at any time through the internet directly to the transfer settlement and form E-business environment. Ecommerce Security is a part of Information Security framework and is basically applied to components affecting e-commerce which includes Computer Security, Data security and other wider sphere of Information Security framework. E-commerce's safety has its awesome degree and is one of the maximum visible security constituents affecting users through ordinary payment interaction inside businesses.

We discuss about various security measures and protocols used for the security of online transaction in which electronic cash flows from buyer to the supplier or merchant. For that there are mainly two challenges the first challenge comes from SET (Secure Electronic Transaction) s is how to use digital envelopes. One part of a digital envelope is the main body of the message. The other part contains that key and is encrypted with the recipient's public encryption key. The two parts may have some common data, possibly hashed, in order to confirm that they are tied together. The second challenging aspect of the SET protocols is the possibility for cardholders and merchants to make public/private key pairs as they want for their electronic credentials.

Here we have introduced a system digital wallet for secure e-payment. System has security at the registration level as well as at transaction time. There is always threat to digital wallet due to hackers. We know that during e-commerce transaction the security threat get increased. So, this research is an attempt to make e-commerce system more secure and prevent the unauthentic operations. System would definitely help in securing e-commerce transaction. There may be two cases of online transaction. One is the situation when users pay for product from his bank account. Another situation is when user pays for product from his digital wallet.

There are some steps in processing online payment or e-payment. We need to follow these steps to keep our money safe.

- The payment procedure is initiated by the applicant. The applicant selects a bank.
- A payment request is sent to the bank that contains an XML message with a redirection URL that points to the government application. In response, the bank opens a session and forwards the user to the given URL.
- The authority's application forwards the applicant on to the online banking application of his bank. After he has been authenticated, the payment transaction is carried out.
- Before the transaction is carried out, the bank checks if there is still a connection open between the bank and the authority.
- After the connection is confirmed by the authority, the bank carries out the money transfer.
- A confirmation message is sent to the authority .
- The authority responds with an acknowledgement message.
- The payment process is finalized and the applicant is referred back to the authority's application.

There are the steps carried out for payment during online procedures.

There are some various methods for e-payment system. Some systems are quite similar, and differ only in some minor details.

- Credit card
- Debit card
- Electronic cash
- Electronic cheque
- Smart card
- Digital wallet (electronic wallet)

This studied have been carried out on E-Payment System. This study focuses on user acceptance e payment system. The study to know the factors which are supporting and which are opposing while implementing a particular e-payment in India. Data used in this study collected basically from the secondary sources. The relevant data was collected based on e-payment system and which e payment type of most suitable.

Questions asked in survey are related to E-Payment system in which given options are Agree, Disagree, strongly disagree, strongly agree, Neutral.

II. Electronic Transactions Vs Payments

A. Electronic Transactions

Electronic transaction protocols deal with the transactions as a whole. E-transaction hence means a business process where money is transferred electronically from one place into another. It could be through internet banking, ATM, from stock exchange trades or just an invoice completion for some service/goods. Nowadays almost all transactions are classed as electronic transactions. An electronic transaction either fails or all of its operations are carried out. If a transaction fails all partially completed results will be undone. All transactions that complete successfully cannot be undone, and the results of these transactions are not lost.

B. Electronic Payments

Electronic payment deals with the actual money transfer. Normally e-payment is done via debit cards, credit cards or direct bank deposits. But there are also other alternative payment methods such as e-wallets and crypto currencies

1) Common Methods of E-Payment:

- Credit or Debit Card Payments - When paying by card, an electronic payment device initiates the online payment transfer. The consumer will have to fill out their card details and have their transfer cleared by their bank.
- Internet Banking - This is done by digitally transferring funds over the internet from one bank account to another. This method tends to be used by smaller businesses or freelance workers.
- Direct Debit - Direct debit transactions transfer funds from a customer's account with the help of a third party.

- E-Wallet - An E-Wallet is a form of prepaid account where the customer's account information is stored electronically. This means you can quickly and seamlessly pay for goods online without having to constantly fill out your information.
- Store Credit Card - A store credit card is a card that has a prepaid monetary value. A more common name for this type of payment is a gift card.

III. Electronic Payment Models

Electronic payment systems can be grouped into three broad classes:

A. Traditional Money Transactions

For secure credit card transactions, a customer's credit card number is encrypted using public key cryptography, so that it can only be read by the merchant. The big advantage of this approach is that the customer does not need to be registered with a network payment service; all that is needed is a credit card number. However, anyone with knowledge of the customer's credit card number can create an order for payment. Also, because payments processed using this approach are processed as standard credit card charges, the costs are quite high so that this method is not suited for payments whose amounts are on the order of cents.

1) Current System: There are several systems that facilitate secured credit card transactions over the Internet

- SET
- PCT
- IKP
- First Virtual's Info Commerce System

B. Digital Currency

Digital currency (digital money, electronic money or electronic currency) is any currency, money, or money-like asset that is primarily managed, stored or exchanged on digital computer systems, especially over the internet. The digital money, an encoded string of digits, can be carried on a smartcard, or stored on a computer disk. One cardholder can make a payment to another without bank involvement, by placing both cards in a 'digital wallet' that moves coins from one card to the other. Types of digital currencies include crypto currency, virtual currency and central bank digital currency.

1) Current System:

- DigiCash
- NetCash
- CAFÉ
- Mondex
- Brands' off-line electronic cash system

C. Credit Debit Payments

Debit and credit card payments are the most common electronic payment methods. For these payments, payment details are entered manually during a checkout process, or the physical card is processed at a point of sale (POS) system. Credit and debit card payments typically will clear into your bank account faster than other electronic payment methods. Electronic credit/debit money can be represented by a digital cheque, just like a normal bank cheque. It includes CMU's NetBill, First Virtual's Info Commerce system, and USC-ISI's Net Cheque system.

1) Current System:

- Millicent
- NetCheque (Neuman and Medvinsky 1995)
- UEPS

D. Size of Transactions

Three classes of typical transactions:

- Tiny value transactions: below $1 (trade-off between convenience of transaction vs. the security or durability of transaction)
- Medium value transactions: between $1 and $1,000 (All the system can perform medium value transactions.)
- Large value transactions: above $1,000 (Require highly secure protocols whose implementation are costly)

IV. Requirements

A. Concerns of Customers and Merchants

Customers and merchants will have an almost common set of wishes and concerns for electronic commerce mechanisms:

- Security.
- Acceptability.
- Convenience.
- Cost
- Privacy
- Durability.

B. Requirements for Financial Institutions

- Immediate Control
- Traceability.
- Control over the spread of encryption mechanisms.

C. Trade_offs

- Privacy versus traceability
- On-line versus off-line
- Hardware versus software
- Transparency versus explicitness

V. Issues and Challenges of E Payment System

Electronic Payments provides greater freedom to individuals in paying their taxes, licenses, fees, fines and purchases at unconventional locations and at any time. But there are some issues and challenges regarding to e payment systems. They are

- Lack of usability:
  Electronic payment system requires large amount of information from end users or make transactions more difficult by using complex elaborated websites interfaces. For example credit card payments through a website are not easiest way to pay as this system requires large amount of personal data and contact details in web form.

- Lack of security:
  Customers have to provide credit card and payment account details and other personal information online. This data is sometimes transmitted in an un-secured way
- Issues with e-Cash:
  The main problem of e-cash is that it is not universally accepted because it is necessary that the commercial establishment accept it as payment method. It should be widely accepted. Another problem is that when we make payment by using e-cash, the client and the salesman have accounts in the same bank which issue e-cash. The payment is not valid in other banks.
- Lack of Trust:
  Electronic payments have a long history of fraud, misuse and low reliability as well as it is new system without established positive reputation
- Users perception regarding acceptance of electronic payment systems:
  The studies shows that users attitudes and human factors are important aspects affecting the success of any information system. These are not successful without acceptance of users.
- Lack of Awareness:
  Even educated people also face problems in making online payments.

A. Solutions to improve the e-payment system quality

- Encryption:
  Most online payment systems use an encryption system to add security to the transmission of personal and payment details. There are various encryption schemes in use to prevent from frauds of online payments.
- Digital Signatures:
  The parties involved in online payments; transactions should use digital signatures in order to ensure authentication of transactions.
- Firewalls:
  A firewall is an integrated collection of security measures designed to prevent unauthorized electronic access to a networked computer system to protect private network and individuals' machines based on firewall policies
- Check Whether the Country is a "High Risk" Country:
  Always require closer inspection for orders that being shipped to an international address. Pay more attention if the card or the shipping address is in an area prone to credit card fraud.
- Call the credit card issuing bank to verify the validity of credit card:
  If online merchants have any suspicions about an order and need to confirm the details of the order, they can call the issuing bank and ask to confirm the general account details. This is to make sure that the card is not stolen.

VI. Strategies for the Security of E-Payments

online payment is not only a type of payment on internet, but also the important basis of internet economic activity. however, due to the openness of the network itself, as well as the constraints of the current network technology, online payment faces with various security threats, and they also raise a variety of security issues in the transaction process, which includes,

- The leakage of customer information
- The transaction fraud problems
- The online payment network system risk

In cause analysis of online payment security issues, we can analyse in sight of man-made cause and natural cause, which are large influence on online payment security factors. The anthropogenic causes include,

- The lack of safety awareness and poor operating habits
- Consecutive malicious activity, Disrupting the market order

And the random causes include

- Imperfect security protocol
- Browser Spy on User Privacy
- What Results in The Derivative Problem is Online Payment Platform Itself?

Countermeasure for online payment security issues is

- The Correct guide customers to problem of cognition, culture their confidence
- Improve laws and regulations, strict the implementation of a system of penalties to keep a norm order
- SET and SSL coexistence and complementary advantages
- The browser company should balance its interests and social responsibility
- Maintain the online payment platform system and risk control
- To strengthen consumer protection by insurance companies

To summarize, all aspects of online payment, such as security protocols, browsers, users, businesses all should improve, being perfect, jointly guarantee the security of the payment.

A. Set Registration Protocols

Everyone normally pay for goods purchased over the Internet bygiving the merchant their credit card details. To prevent credit card information from unwanted people from stealing the card number, the message undergoes a session of the secure sockets layer (SSL) protocol. For that mainly two types of registrations are there the first one is registration of cardholder in that the cardholder initiates registrations requests, after the approval from the authorities he then requests the registration form and the authorities responds accordingly. After that cardholder request for certificates, and response to that request is obtained by receiving the certificate. And the second one is registration of merchant in that the merchant requests the registration forms, after the approval from the merchants request for certificates, and response to that request is obtained by receiving the certificate.

B. Sensitive Data Transfer Secure

Sensitive data transfer secure (SDTS) model is converting the secure information into barcodes first, this information converting can be done by pixel manipulations and then convert the barcode image into byte arrays they may be even odd byte array. Finally, it encrypts the bytes using standard RSA algorithm. The information to be processed is not sent over the network in unsecured manner, a security key corresponding to the information is picked from database table to secure the real time data or information from being exposed on the network or to the people who are not intended to view it. Then secure/secret key before exposing to the network is processed under various layers and converting into unreadable form. A quick response barcode is created for the key and then it is splitted into two identical barcode like images.

C. Onsite Transaction Procedure

This is the scenario where the cardholder is physically present at the merchant's site or shop and gives his smartcard at the merchant's terminal after selecting the goods and services to purchase. Customer/cardholder gives his smartcard at the merchant's terminal and then brings the smartcard near the smartcard reader. The Credit card details are encrypted using public key. The cardholder sends a onetime password to the mobile device of the cardholder through the cellular network. Cardholder enters the one-time password obtained on his mobile device on to the merchant's terminal. After finally authorizing the cardholder, it passes the payment details to the issuing bank. After verifying the payment information issuing bank transfers the respective funds to the merchant's account.

D. Biometric Authentication

Biometrics operation is very common application for identification across the world. Biometric authentication technologies such as face, finger, hand, iris, and speaker recognition are used largely. A biometric system is mainly a pattern recognition system that operates by taking biometric data from an person, extracting a feature from the data, and comparing extracted features with data stored in the database. Biometric system operates in two modes verification mode and identification mode.

In the verification mode, the system validates a person's identity by comparing the captured biometric data with data stored in system database. The aim is to prevent multiple people from using the same identity and thus achieving security of the system. In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Therefore, the system conducts a one-to-many comparison.

E. Secured Fingerprint Payment System

A payment application would be installed onto a android device, for authentication fingerprint is taken at runtime. The finger print template would be captured by the phone and compared against a stored template on a database server. The fingerprint template is encrypted by using the RSA algorithms or any other encryption algorithm and sends it to the Bank. Fingerprint is used for the login purpose for the bank application on mobile. Mobile will act as a client in this system and the bank website will act as a server in this system. Once fingerprint is taken as a login by mobile device, it sent to the server for matching as request, and server send the reply message. If it is matching then only login will be successful and user can do the transaction otherwise user will not be given access to online electronic transaction and this security of system will be achieved.

VII. Implementation of Security – Digital Wallet

Digital wallet broach to a digital device that permits someone to make digital transactions. A person's bank account can also be linked to digital wallet. They could also have their driver's license, health card, loyalty card(s) and other id documents saved on phone. The credentials might be surpassed to a merchant's terminal wirelessly thru near discipline conversation (NFC). Increasingly more, digital wallets are being made not only for primary economic transactions however to also authenticate holder's credentials. Types 0f securities used in this system are

- Securing the user data at the time of storing in database using cryptographic algorithm
- Allow user to access data when user login to allow him to access his own account.
- Securing the transaction using a pattern lock.
- Securing digital wallet by correct logging and pattern lock.
- One time password security during transaction.

As the output, the system includes a signup form, Login panel, Add product form, place order form and a transaction module. When the admin clicks on add Product, he's allowed to add products to his cart. The homepage of the application suggests options for product list, logging in and signing up Views. The signup form is data associated with person is entered. Place order form is used to take input from user in order to place order. The person who has logged in could login from login panel. The output windows include View product list, View Invoice, View list of orders as well as transaction are listed with the help of the buttons on Admin panel and accessing all data from Remote Database Server. The size of traditional ecommerce application is about 30 megabytes but proposed application size is about 1 megabyte. And the application is more accessible and lighter weighted because it uses swings rather than GUI. The performance rate of the system is very high.

VIII. Conclusion

Technology has arguably made our lives easier. One of the technological innovations in banking, finance and commerce is the Electronic Payments. Electronic Payments (e-payments) refers to the technological breakthrough that enables us to perform financial transactions electronically, thus avoiding long lines and other hassles. Electronic Payments provides greater freedom to individuals in paying their taxes, licenses, fees, fines and purchases at unconventional locations and at whichever time of the day, 365 days of the year. After analysis and comparison of various modes of electronic payment systems, it is revealed that it is quite difficult, if not impossible, to suggest that which payment system is best. Some systems are quite similar, and differ only in some minor details. Thus, there are number of factors which affect the usage of e-commerce payment systems. Among all these user bases is most important. Added to this, success of e-commerce payment systems also depends on consumer preferences, ease of use, cost, industry agreement, authorization, security, authentication, non-refutability, accessibility and reliability and anonymity and public policy. The Reliable and Cashless payment system offers immunity against theft of paper and e-money, and adopting e-payment solutions or systems for different reasons. In addition to cost reduction, reference was made to a number of other benefits, including improved customer service, improved working capital, increased operational efficiencies and cycle times, processing efficiencies and enhanced compliance to organizational policies and procedures. These opportunities e-payment operation increases different levels of risks for marketing. More than ten Years of Internet marketing research have yielded a set of important findings. Based on our review of these findings, it is clear that the Internet is playing a more and more important role in the field of epayment. peoples are becoming aware of the need to measure the collaborative effects of e-payment The study reveals that the peoples were not aware and educated. They have not any knowledge of epayment. The study is based on survey. The respondent has to answer the questions on their own. Some people satisfy with our views. But some peoples are not satisfying with us. This study states that Online e-payment provides greater reach to customers. Feedback can be obtained easily as internet is virtual in nature. Customer loyalty can be gain. Personal attention can be given by bank to customer also quality service can be served.

Also in this research is an attempt to make e-commerce system more secure and prevent the unauthentic operations. System would definitely help in securing e-commerce transaction. There may be two cases of online transaction. One is the situation when users pay for product from his bank account. Another situation is when user pays for product from his digital wallet. There is always risk to such situation as banking sites have their own security mechanisms but the security of users amount in digital wallet is responsible is provided by its makers. Here we have made such digital wallet and secure it using pattern lock and one time password facility. To summarize, all aspects of online payment, such as security protocols, browsers, users, businesses all should improve, being perfect, jointly guarantee the security of the payment.

References

[1] Kaur Manjot, E-Commerce, Kalyani Publcation, New Delhi(2012).

[2] Karamjeet Kaur , Dr. Ashutosh Pathak " E-Payment System on E-Commerce in India ", ISSN : 2248-9622, Vol. 5, Issue 2, (Part -1) February 2015, pp.79-87

[3] Nikhil Khandare, Dr. B. B. Meshram " Security of Online Electronic Transactions", ISSN: 2278-0181, Vol. 2 Issue 10, October - 2013

[4] Rachna, Priyanka Singh "Issues and Challenges of Electronic Payment Systems", ISSN: 2320- 0901, Vol. 2, Issue 9, December 2013

[5] C hen Zhang, Shijie Jiang, Bin Huang," Strategies for The Security of Online Payments in E-commerce", Proceedings of the 2012 2nd International Conference on Computer and Information Application (ICCIA 2012)

[6] Paul J.M. Havinga, Gerard J.M. Smit, Arne Helme "Survey Of Electronic Payment Methods And Systems"

[7] Purva Goyal," Implementation on Enhancing Security of E-commerce Site for Digital Marketing", 2018 IJSRCSEIT — Volume 3 — Issue 1 — ISSN : 2456-3307

[8] Bhasker, Bharat (2013). Electronic Commerce, Framework, Technologies and Applications. McGraw Hill Education (India) Private Limited., p.9.2-9.16.

[9] Denis Abrazhevich." "Electronic Payment Systems:a User-Centered Perspective and Interaction Design", Proefschrift. -ISBN 90-386-1948-0NUR 788, 2004.

[10] Prof. Sana Khan, Ms. Shreya Jain " A Study on Usage of ePayments for Sustainable Growth of Online Business", e-ISSN: 2278-487X, p-ISSN: 2319-7668 PP 74-81

[11] AsafShabtai,YuvalFledel, Uri Kanonov, Yuval Elovici, ShlomiDolev(2010), —Google Android: A Comprehensive Security Assessment.‖ IEEEsecurity and Privacy.

[12] D. K. Gangeshwer, "E-Commerce or Internet Marketing: A Business Review from Indian Context", International Journal of u- and eService, Science and Technology Vol.6, No.6 (2013).

[13] Anderson R.J. "UEPS - a second generation electronic wallet", Computer Security - ESORICS 92, Springer LNCS v 648, pp 411-418.

[14] Kuldeep Kaur1, Dr. Ashutosh Pathak2, "ECommerce Privacy and Security System" Kuldeep Kaur Int. Journal of Engineering Research and Applications www.ijera.com ISSN : 2248-9622, Vol. 5, Issue 5, might 2015.

[15] Revathi C, Shanthi K, Saranya A.R "A Study on E-Commerce Security Issues" International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 12, December 2015.

[16] Niranjanamurthy M "E-commerce: Recommended Online Payment Method – PayPal" IJCSMC, Vol. 3, Issue. 7, July 2014.

[17] A N R Hami & A Y Cheng "A risk perception analysis on the use of electronic payment systems by young adult order" 6(8.4) 6-7 2002

[18] P Aigbe & J Akpojaro "Analysis of security issues in electronic payment systems" International journal of computer applications 108(10), 2014

[19] S k Antwi, K Hamza & S W Bavoh "Examining the effectiveness of electronic payment system in Ghana: the case of e-ZWICH in the Tamale metropolis" Research Journal of Finance and Accounting,6(2) 163-177,2015

[20] C K Ayo & W I Ukpere "Design of a secure unified e-payment system in Nigeria: A Case Study" African Journal of Business Management 4(9),1753-1760,2010

[21] Z Bezovski 2016 The Future of the Mobile Payment as Electronic Payment System (European Journal of Business and Management) 8(8) 127-132

A Perspective Approach on the Cognitive Aspects of Human Computer Interaction

Hridhya K B,
Sruthy Krishna P U,
and Swetha M S
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Sajay K R
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: sajay.k.r@vidyaacademy.ac.in)

***Abstract*—The intention of this paper is to provide an overview on the subject of Human-Computer Interaction. The overview includes basic definitions, terminologies, previous projects, current projects, technologies and applications. This paper also covers the application level of HCI in smart environments. It also offers a comprehensive explanation about the application level of HCI that is affective computing, its background, applications, effects and issues. Ubiquitous Computing is new computer paradigm with seamless integration of hundreds and thousands of self-communicating small scale computers and intelligent devices into the user environment and daily life activities. Through this paper we are creating awareness among the people about this rising field of HCI.**

***Index Terms*—Human-Computer Interaction(HCI); Multimodal HCI; Smart environments; Design of HCI; Natural user interfaces; Assistive technologies; Mobile computing; Ubiquitous Computing; explicit Human Computer Interactions (eHCI); implicit Human Computer Interactions (iHCI); Context aware Systems; Natural interfaces; Multimodal interfaces; Affective computing; Affective HCI; Affective state; Human Centered Computing(HCC); Psychophysiology; Affective intervention.**

I. Introduction

Human-computer interaction (HCI) is a multidisciplinary subject that focuses on computer design and user experience. It brings together expertise from computer science, cognitive psychology, behavioural science, and design to understand and facilitate better interactions between users and machines. HCI involves the study, planning, and design of the interaction between people (users) and computers. The growth in Human-Computer Interaction (HCI) field has not only been in quality of interaction, it has also experienced different branching in its history. Instead of designing regular interfaces, the different research branches have had different focus on the concepts of multimodality rather than unimodality, intelligent adaptive interfaces rather than command/action-based ones, and finally active rather than passive interfaces. Affective computing facilitates more intuitive, natural computer interfaces by enabling the communication of the user's emotional state. Nowadays, the importance of communication between human and computer is inevitable. So, the computer designing must be user friendly which enables it to communicate easily with human and its surroundings by analyzing human natural way of interaction.

II. Human-Computer Interaction: Definition, Terminology

Sometimes called as Man-Machine Interaction or Interfacing, concept of Human-Computer Interaction/Interfacing (HCI) was automatically represented with the emerging of computer, or more generally machine, itself. HCI is a design that should produce a fit between the user, the machine and the required services in order to achieve a certain performance both in quality and optimality of the services. Determining what makes a certain HCI design good is mostly subjective and context dependant.

III. Design of HCI

In design of HCI, the degree of activity that involves a user with a machine should be thoroughly thought. The user activity has three different levels: physical, cognitive, and affective. The physical aspect determines the mechanics of interaction between human and computer while the cognitive aspect deals with ways that users can understand the system and interact with it. The effective aspect is a more recent issue and it tries not only to make the interaction a pleasurable experience for the user but also to affect the user in a way that make user continue to use the machine by changing attitudes and emotions toward the user.

A. Previous projects

- Hand Gesture Recognition using Data-Glove
- Human action recognition in video stream

B. Current projects

- Image Forgery Detection
- Advance high level Content Based Image Retrieval system for Logo and Trademark Distinctivity
- Facial expression recognition from Audio-Visual information using Wavelet and Curvelet transform

- Computer Vision based Virtual Keyboard (Projected or Generic)
- Text Detection and Localization in Natural Scene Images

IV. Existing HCI Technologies

The existing physical technologies for HCI basically can be categorized by the relative human sense that the device is designed for. These devices are basically relying on three human senses: vision, audition, and touch. Input devices that rely on vision are the most used kind and are commonly either switch-based or pointing devices. The output devices can be any kind of visual display or printing device. The devices that rely on audition are more advance devices that usually need some kind of speech recognition. Output auditory devices are however easier to create. The most difficult and costly devices to build are haptic devices. Haptic devices are generally made for virtual reality or disability assistive applications.

Fig. 1. Canesta virtual keyboard

V. Recent Advances in HCI

Recent directions and advances of research in HCI are intelligent and adaptive interfaces and ubiquitous computing.

A. Intelligent and Adaptive HCI

One important factor in new generation of interfaces is to differentiate between using intelligence in the making of the interface (Intelligent HCI) or in the way that the interface interacts with users (Adaptive HCI). Intelligent HCI designs are interfaces that incorporate at least some kind of intelligence in perception from and/or response to users. Adaptive HCI designs, on the other hand, may not use intelligence in the creation of interface but use it in the way they continue to interact with users. Example that uses both intelligent and adaptive interface is a PDA or a tablet PC that has the handwriting recognition ability and it can adapt to the handwriting of the logged in user.

B. Ubiquitous Computing and Ambient Intelligence

The latest research in HCI field is unmistakably ubiquitous computing (Ubicomp). The term which often used interchangeably by ambient intelligence and pervasive computing, refers to the ultimate methods of human-computer interaction that is the deletion of a desktop and embedding of the computer in the environment so that it becomes invisible to humans while surrounding them everywhere hence the term ambient.

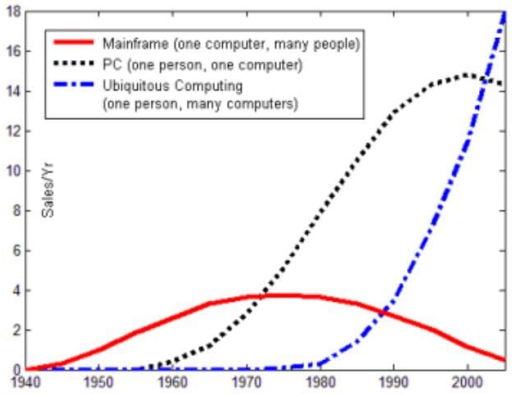


Fig. 2. Major trends in computing

VI. Five Future Goals of HCI

- Communication all-around
- High-functionality systems
- Mass availability of computer graphics
- Large and thin displays
- Information utilities

VII. HCI Systems Architecture

Most important factor of a HCI design is its configuration. Architecture of a HCI system shows what the inputs and outputs are and how they work together.

A. Unimodal HCI Systems

An interface mainly relies on number and diversity of its inputs and outputs which are communication channels that enable users to interact with computer via this interface. Each of the different independent single channels is called a modality. A system that is based on only one modality is called unimodal. Based on the nature of different modalities, they can be divided into three categories:

1) Visual-Based: Different aspects of human responses can be recognized as a visual signal. Some of the main research areas in this section are Facial Expression Analysis, Body Movement Tracking (Large-scale), Gesture Recognition, Gaze Detection (Eyes Movement Tracking).

2) Audio-Based: This area deals with information acquired by different audio signals. Research areas in this section are Speech Recognition, Speaker Recognition, Auditory Emotion Analysis, Human-Made Noise/Sign Detections (Gasp, Sigh, Laugh, Cry, etc.), Musical Interaction.

3) Sensor-Based: This section is a combination of variety of areas with a wide range of applications. These sensors are Pen-Based Interaction, Mouse and Keyboard, Joysticks, Motion Tracking Sensors and Digitizers, Haptic Sensors, Pressure Sensors, Taste/Smell Sensors.

B. Multimodal HCI Systems

The term multimodal refers to combination of multiple modalities. In MMHCI systems, these modalities mostly refer to the ways that the system responds to the inputs, i.e., communication channels. The definition of these channels is inherited from human types of communication which are basically his senses. Therefore, a multimodal interface acts as a facilitator of human-computer interaction via two or more modes of input that go beyond the traditional keyboard and mouse. One of the most commonly supported combinations of input methods is that of gesture and speech.

Applications of Multimodal HCI:

- Smart Video Conferencing
- Intelligent Homes/Offices
- Driver Monitoring
- Intelligent Games
- E-Commerce
- Medicine
- Helping People with Disabilities

VIII. Human Computer Interaction in Ubiquitous Computing

Ubiquitous applications interfaces should be adaptable to their environment and behave according to the situation. They should adapt according to context and do not require explicit interaction from the user. The main objective of the Ubiquitous computing is to allow user to focus on their tasks rather than on technology.

A. HCI Requirements for Ubiquitous Computing

Core characteristics of HCI in Ubiquitous systems are

- HCI in ubiquitous system would be hidden and implicit rather than traditional explicit interactions of user with the system
- Ubiquitous system environment should allow the user to focus on the tasks rather than on the technology
- HCI in ubiquitous computing involves adaptive to situation and support for multimodal of interaction with the user. This requires some sort of intelligence work integrated with interaction paradigm
- Human interaction pattern in Ubiquitous environment would multimodal of interaction support
- With advance in technology the interaction in ubiquitous computing tends to be supported with natural interactions like interaction of human to human

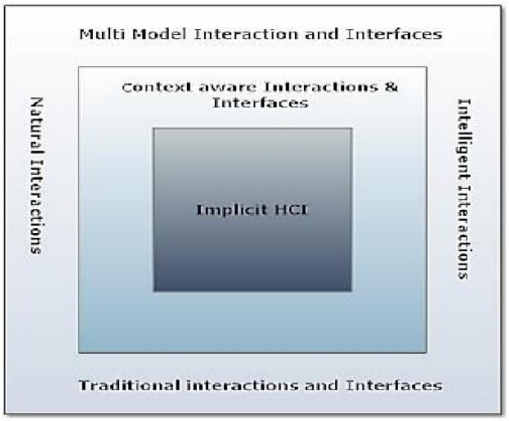


Fig. 3. Ubiquitous computing

B. HCI Framework for Ubiquitous Computing

All these requirements of HCI in ubiquitous computing require an enhanced model of interaction with clear defined goals . The agenda of ubiquitous computing require implicit Human Computer Interactions (iHCI) with the help of context aware interactions and natural, intelligent multimodal interfaces. Traditional interactions theme can also be utilized when they can prove to be useful in ubiquitous computing environment.

C. Implicit HCI

The system should support the user tasks without much intervention and with minimal distraction of user from their core activities. The system tends to disappear and become calm because human interactions with this system tend to implicit rather than explicit. Due to effective use of implicit mode of interaction the calm or disappearing computers would be reality in very near future.

D. Context Aware Interactions

Ubiquitous application has to be context ware system. Previous research on the issue of HCI in Ubiquitous computing has posed the agenda of context ware applications in order to be proactive. Context aware systems are system that are aware of their situation and adopt accordingly. Context aware systems are aware of their environment, situation or context so that system can adapt to the situation and provide services according to the changed situation or context. Context information's depends upon the type of services and application like location, temperature, light etc.

E. Intelligent, Natural and Multimodal Interfaces

Intelligent user interfaces aim to enhance the efficiency, naturalness, and effectiveness of HCI by using artificial intelligence techniques. Single independent channel of interaction is called modality. This is the channel through which information

transferred between system and user. A system that is based on single channel is called unimodal whereas system with multiple channels is called multimodal. The most common system based on different modalities is visual, audio and sensors based Due to multiple focus of HCI research there is also trend in multimodality interfaces from traditional interfaces. Along with shift from traditional to multimodality there is also trend of intelligent, context aware, adaptive interface with implicit human computer interactions rather than explicit interaction pattern through command line and GUI interfaces.

IX. Human Computer Interaction in Smart Environments

Sensors continue to rapidly evolve, becoming increasingly smaller, cheaper, more accurate, more reliable, efficient, responsive and also increasing communication capability. These key factors, as well as the availability of new technologies, are contributing to the growth of the market of consumer electronics sensors, thus reducing their costs. This scenario fosters the embedding of sensors in the everyday objects of our lives. The integration of ubiquitous sensing and networking technologies enables the development of new applications in a wide variety of domains such as, among others, smart homes, e-health, and intelligent transport systems.

Smart systems often make use of camera sensors to capture images and video of the user interacting with the system. Digital colour cameras can be used as sensing devices for inferring human's hands position, poses and gestures, to be translated into suitable commands for the control of virtually every kind of digital system.

A lot of challenges should be faced to guarantee a robust tracking system. A solution to the problem of hand tracking failures by leveraging on context information such as secondary motions of the body or hand-kept objects. A new interaction setup combining glass-free 3D visualization and gesture-tracking technology was proposed to provide natural representation and navigation of 3D scenes. Interaction with 3D scenes is enabled by the analysis of contactless hand and finger gestures.

A system was proposed for inferring the pinch-to-zoom gesture in real-time, based on surface EMG (Electromyography) signals. Pinch-to-zoom is a gesture commonly used to control the size of images or webpages on smart devices, and is based on the distance between the thumb and the index fingers. A new method for improving face recognition systems is based on the improvement of Sparse Representation-based Classification algorithms (SRC), which recently proved to be robust and effective in the recognition process. A presented system for assessing eye fatigue caused by 3D displays by analysing multimodal measurements collects electroencephalograph (EEG) signals, facial temperature (FT), information on eye blinking rate (BR), as well as subjective evaluations (SE) before and after a 30 min 3D movie watch and evaluates correlation values.

A couple of works are related to the exploitation of sensors in smart vehicles, which analyse data concerning driver's vital parameters. In particular, a system was proposed to jointly use palmar perspiration and masseter electromyography sensors to assess truck drivers' mental stress. In another work EEG sensors are used to estimate eye closure degree, in order to detect driver drowsiness.

Some works deal with the exploitation of sensors for improving human-robot interaction (HRI). Room furniture has been equipped with Radio-frequency identification (RFID) readers and load cells in order to accurately identify the position of objects. A system is proposed to rely on colour camera sensors to build interactive interfaces for laser guided robots. These interfaces exploit laser spot tracking methods to extract different shapes from laser traces, which are associated to robot commands.

Human-computer interaction in aircraft navigation systems

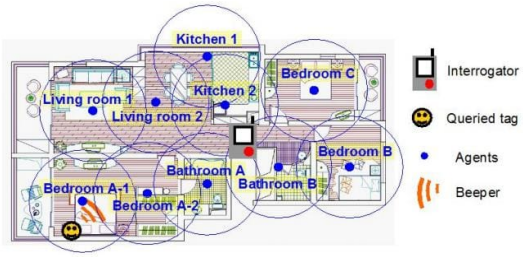


Fig. 4. RFID-Based Object Locators

plays a key role for enhancing situation awareness and supporting decision-making processes. In this sense, a step towards the development of smart air traffic systems where inputs coming from sensors provide the pilot with controller indications as well as surrounding traffic situation through the cockpit.

X. Background of Affective Computing

HCC broadly describes the methodology that would be applied to any field that uses computers in any form where users directly interact with them. Thus, HCC aims to integrate human sciences into the existing body of computer science and HCI knowledge with a human focus throughout the lifecycle. HCC is said to incorporate social and cognitive sciences more closely than traditional HCI. HCC principles of cognitive and affective design are important when developing usable software and systems. A successful user interface would have an understanding of what emotions are, how they can be identified and what the implications of various emotional patterns are for a given interaction situation.

The interaction between affect and cognition is bi-directional, thus the underlying affective state of the individual will also influence the outcome of various cognitive processes. Emotion has an impact on the speed at which information is processed. Emotion also has a relation to motivation in that evaluations or feelings regarding the current situation will largely determine the action that is taken in response. Memory is also impacted by emotional state.

Although a computer will not actually experience emotions in

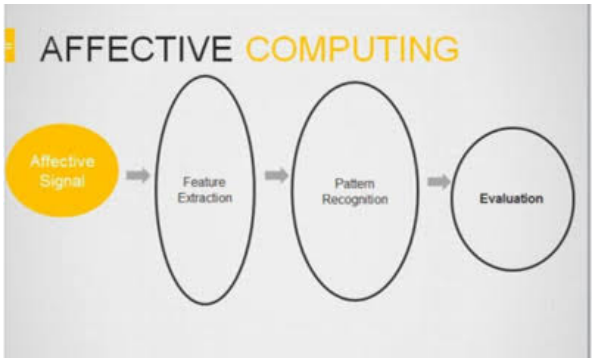


Fig. 5. Affective computing

the same way that a human would, the quality of interaction has been shown to improve even if the system appears to do so. The results of some experiments demonstrated that if the computer system provided users with the ability to vent their frustrations , users continued to interact with the frustrating system significantly longer than if no affect-support was provided. Empathic agents have also been successfully used in software to improve usability. Other certain findings from experiments indicated that a character demonstrating empathy may decrease users' levels of experienced stress.
Mehrabian (1981) is often cited for his 7%-38%- 55% rule of non-verbal communication which simply states that in human communication, 7% of the message is communicated by the words, 38% by tone of voice and the remaining 55% by body language. Methods for inferring affective state are numerous but may be broadly categorized into a few areas, including self-report measures, observable traits and psychophysiology.

A. Applications of Affective Computing

Affective computing applications have potential uses in practically any situation where a human-computer interaction is taking place. Measuring the stress or difficulty caused by a system may also allow developers to pinpoint problems, or simply allow the system to be improved by being able to respond in a more natural and realistic way. Such technologies have been successfully implemented in very diverse environments. These include robotic personas, learning companions , or wearable computers.
Sociable humanoid robots present a novel and under explored area of human-machine interaction. Traditionally robots have been utilized for functional roles such as automation it can also be used for any number of applications, including domestic use.
Humanoid robots may express affective states in a number of ways. The robot possesses a computation model of emotion, in which affective state is sensed, used as part of the internal decision (action) making strategy, and communicated to the human.
Education is an area in which affective computing applications have shown promise, partly due to the increasing reliance on online and computer mediated learning and teaching strategies. Intelligent tutoring systems incorporating an emotional or affective model are known as affective tutoring systems. Eg.AutoTutor is an intelligent tutoring system that interacts with learners using natural language and helps them to construct explanations in simulation environments. These systems have also been shown to be effective and result in increased learning, however are still not as effective as a one-to-one human tutor.
Wearable computers provide a rich and diverse ground for evaluating and implementing affective technologies. Existing devices that are in close contact with the user may also have potential to be used as affect sensing devices, for example a mouse may sense the user's stress levels, or a car steering wheel may sense when the user is falling asleep or identify lapses in attention.

B. The effects of Affective Interventions in Human Computer Interaction

The field of human–computer interaction has recently been widened by a new research field of affective computing. Picard defined affective computing as computing that relates to, arises from, or deliberately influences emotions. The growing interest in affective computing originates from findings in psychology and psychophysiology of human emotions, which highlight the importance of emotion in human behavior.
Emotion can be defined as discrete categories or as a space of different dimensions. Ekman defined six basic emotions: anger, disgust, fear, joy, sadness, and surprise. On the other hand, according to the dimensional emotions framework, emotion can be defined as a set of dimensions, which together define the emotional space. Out of the original three dimensions (valence, arousal, and dominance), valence and arousal are the two most commonly used dimensions. Many studies have shown that various changes in physiological activity are integrally related to emotional responses. By measuring the activations of lip corner, eye brow like muscles is possible to obtain information about the subject's affective state. In current emotion research, it is in most cases imperative to monitor human physiology while studying emotions. Other proposed physiological measures have included, like pupil size measurement.
There are several possible techniques for the emotional stimulation of the user. In addition to textual or other visual emotional messages, auditory stimulation offers a promising alternative. There is evidence that different emotional reactions can be evoked using different auditory stimuli. Emotionally provocative sounds can be used to evoke different autonomic nervous system responses and subjective experiences. Facial EMG responses were measured during and after positive and negative worded affective interventions. Task performance in problem solving was analyzed following these interventions. Facial EMG activity and task performance were also measured in relation to a no intervention condition.
In human–computer interaction, the transition from affective feedback to affective interventions seems to open new possi-

bilities. In addition to using affective interventions to alleviate the effects of system problems, there are application fields, which could benefit from affective interventions. Especially in fields in which motivation is a crucial factor, such as computer-assisted learning, affective interventions could be assumed to have beneficial effects.

Clearly, a great deal of further research is needed to understand the effects of affective interventions in different situations. The contexts in which the interventions are given also range from feedback related to different user tasks to interventions not directly related to the user's task like in this experiment. In addition, affective interventions could be used as computer responses to significant changes in the user's physiology. Following the knowledge on how to recognize emotional states from affective signals, the question of how to react to the changes in the user's affective state becomes important. Physiological signals can be monitored in real time by the computer, and the computer can also monitor the physiological changes caused by its affective interventions, and adjust its behavior accordingly. Consequently, this kind of advanced adaptive technology could even show emotional intelligence similar to that of humans. Thus, based on the current experiment, we can seriously encourage further studies on the effects of affective interventions in human–computer interaction.

C. Issues

Many of the implementations are disparate and "one-off" in nature, it is potentially difficult to transfer findings from one particular domain to a new application. Affective computing applications are often built in the same way as more traditional applications, with the affective functionality inserted into the program architecture wherever the developer considers it appropriate.

Implementation environments which may have different outputs, manufacturers and operating requirements. Due to the diverse nature of implementation, and the many methods by which affective information may be gathered, affective computer interface com ponents often need to deal with many different types of data which presents a complex signal processing task.

XI. Conclusion

Human-Computer Interaction is an important part of systems design. Quality of system depends on how it is represented and used by users. Therefore, enormous amount of attention has been paid to better designs of HCI. HCI involves the study, planning, and design of the interaction between people (users) and computers. It is often regarded as the intersection of computer science, behavioral sciences, design and several other fields of study. The new direction of research is to replace common regular methods of interaction with intelligent, adaptive, multimodal, natural methods. Ambient intelligence or ubiquitous computing which is called the Third Wave is trying to embed the technology into the environment so to make it more natural and invisible at the same time. Ubiquitous computing has certain requirements and challenges in context of Human Computer Interactions (HCI). The spirit of Ubiquitous computing demands that there should be implicit Human Computer Interactions (HCI) in order to allow user to focus on their task with minimum distractions. As a whole, the results of the current study showed that as compared to no intervention condition, different physiological, behavioral, and experiential user reactions could be evoked by using affective interventions. The results showed that significant changes in physiology and subjective experiences could be evoked by altering the emotional wording of the intervention. As such, affective computing holds great promise for improving human-computer interaction. This article has provided an insight into the field of affective computing: covering the origins of the field, the underlying elements of affective computing, and highlighting issues in the field. Detailed examples of a diverse range of affective computing applications are provided and future directions have been identified.

References

[1] R Harish, Shahbaz Ali khan, Shokat Ali, Rajat ,Vaibhav Jain, "Human Computer Interaction-A brief study," International Journal of Management,IT and Engineering, VOL. 3, ISSN 7, pp 390-401, July, 2013.

[2] Rab Nawaz Bashir, Salman Qadri, Rana Muhammad Saleem, Muhammad Naeem, "Human Computer Interaction(HCI) in Ubiquitous Computing," International Journal of Innovation and Applied Studies, Vol. 9, No.2, ISSN 2028-9324, November, 2014.

[3] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, Mo Nours Arab, "Human-Computer Interaction: Overview on State of the Art," INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS, VOL. 1, NO. 1, March, 2008.

[4] Gianluca Paravati, Valentina Gatteschi, "Human-Computer Interaction In Smart Environments," Sensors, Vol. 15, ISSN 1424-8220, pp 19487-19494, August, 2015.

[5] Timo Partala, Veillo Surakka, "The effects of affective interventions in human-computer interaction," Interacting with Computers, VOL. 16, pp 295-309, February, 2004.

[6] Nik Thompson,Tanya McGill, "Affective Human-Computer Interaction," Encyclopedia of Information Science and Technology, Third Edition, pp 3712-3720, December, 2014.

[7] B. Shneiderman," Designing the User Interface," Strategies for Effective Human-Computer Interaction(3rd edition), Addison Wesley Longman, Reading 1998.

[8] Mohamed Hashem, Nagwa Badr, Hanaa Talha, Raghda Fouad, "Exploring a Hybrid of Geospatial Semantic Information in Ubiquitous Computing Environment," International Journal of Computer Science Issues, Vol. 8, No.6, November, 2011.

[9] Aula A, Surakka V, "Auditory Emotional Feedback Facilitates Human-Computer Interaction," Proceedings of HCI ,pp.337-349, March, 2002.

[10] Allanson, Fairclough, S H, "A research agenda for physiological computing," Interacting with computers, vol.16, pp 857-878, August, 2004.

[11] J.Y. Chai, P.Hong, M.X. Zhou, "A probabilistic approach to reference resolution in multimodal user interfaces," Proceedings of the 9th international conference on Intelligent User Interfaces, pp 70-77, December, 2004.

[12] B A Myers, "A brief history of human -computer interaction technology," ACM Interactions, Vol. 5, pp 44-54, 1998.

[13] O.Khatib, O.Brock, K.S. Chang, D. Ruspini, L. Sentis, S. Viji, "Human-centered robotics and interactive haptic simulation," International journal of robotics research, Vol. 23, pp 167-178, 2004.

An Overview of Cyber Security: Issues, Attacks, Goals and Models

Afra P L, Hansa A M and Naina Vincent
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Manesh D
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: manesh.d@vidyaacademy.ac.in)

***Abstract*—Cyber security has been used interchangeably for information security, where later considers the role of the human in the security process while former consider this as an additional dimension and also, focus person has a potential target.To address the issue of cyber security attacks, various frameworks and models have been developed. In the context of cyber security, social engineering is the practice of taking advantage of human weaknesses through manipulation to accomplish a malicious goal. Cyber threat intelligence (CTI) provides an organization the capability to identify, gather, analyze threats and the associated risks so that CSC organization can forecast the existing and future threat trends and manage the cybersecurity risk in a proactive manner.The purpose of this paper is to give a review on cyber security, its goals, impacts, models and issues.**

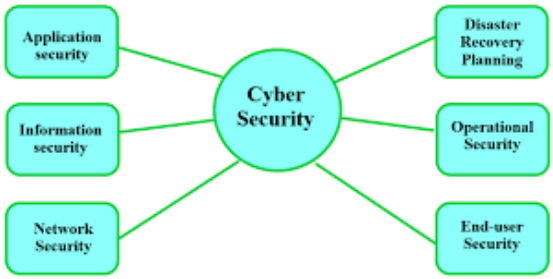
***Index Terms*—Cyber security,Threats,cyber security awareness,social engineering,cyber threat intelligence,ontologies**

I. Introduction

Cyber security means integration of data, security, storage and transfer of data through electronic or other modes. Cyber security indicates three important factors. The methods of protecting Information Technology (IT), the data itself, the data being processed and transmitted together with physical and virtual setup, the level of protection obtained by applying such measures and the professional aspects associated. Lack of cyber security is the main reason behind rapid increase of cyber crimes.Cyber attack commonly known as a computer network attack (CNA) is the deliberate exploitation of computer systems, companies and networks dependent. Here are some Cyber Security concerns associations may confront:

A. Unprecedented Attacks

Organisations uses multiple gadgets to store their confidential data. They get the security for their data by changing their working process,rules and regulations and increasing the number of gadgets.But the hackers will hack the information by increasing the number of gadgets.So the oraganisations want to change the usage of gadgets,then only their datas will be secured.



B. Cyber Espionage

Cloud meaning- storing and accessing data and programs over internet instead of computer hardware . At present most of the companies have opted to store their confidential information in the cloud. Since it has lot of risk involved,the data's can be hacked out easily.The security check is not the issue with cloud , it the way you configure cloud computing and it involves the absence of experience in using the platform.

C. Data Theft

Hackers have kept the small and medium enterprises to be their easy targets.Those small and medium type enterprises have a lack of security management in handling confidential datas . As they dont have large and robust security protocols to prevent the theft. Understanding the the technology is required and the proper implementation of security check on the datas stored should be maintained inorder to avoid data theft.

II. Major Cyber Attacks

The classification of the threat picture correlates with the complexity and severity to carry out various attacks. For this purpose, the threat taxonomy has been subdivided into three attack categories: low, medium and high. Low-level attacks may be executed by widespread available equipment with minimal effort and marginal knowledge of the ATC system. Whereas, medium-level and high-level attacks require

a particular transmitter and specific knowledge to jam or modify ADS-B messages.

- Eavesdropping
  This type of attack is called a passive attack which is the most straightforward form of listening into unencrypted broadcast transmissions. As ADS-B provides unauthenticated information over an unencrypted data link, it can be captured in real-time. Eavesdropping may form the basis for a number of more sophisticated active attacks. It should be also noted that eavesdropping is almost impossible to prevent and to detect in case of an unencrypted data transmission. In terms of the threat taxonomy, eavesdropping is classified as low-level attack, due to the reason that everyone can easily use flight tracking services to track air traffic in real time.

- Jamming
  This kind of attack is used to attack a single node (e.g., aircraft, ground infrastructure or both) or an area with multiple users, which are not able to receive ADS-B messages anymore. Since jamming is a common problem in wireless communication, the effects in aviation are far more severe. In particular, the difficulty lies in the wide-open spaces that cannot be controlled, as well as the criticality of the transmitted data. Jamming attacks are more difficult due to the higher transmission power and the rotating antenna. Jamming attacks are classified as low to medium-level attacks regarding to the threat taxonomy. The consequences can range from minor to high flight safety risks.

- Injecting
  It is possible to inject malicious ADS-B messages into the ATC system. Once, there is no authentication scheme implemented at the data link layer of ADS-B, an attacker does not need a special transmitter which is able to produce correct and formatted ADS-B messages.Injections are classified as medium to high-level attacks, which highly affect the flight safety.
  itemDeleting
  Valid ADS-B messages can be deleted by using destructive or constructive interference. The key for deleting ADS-B messages to achieve an aircraft disappearance is established by injecting ADS-B messages into the ATC system.Destructive interference: In theory, an inverse signal of a valid transmitter is broadcasted. Based on the resulting superposition, the signal should be deleted or at least highly attenuated which may results in an aircraft disappearance. Deleting valid ADS-B messages is classified as high-level attacks, which highly affect the flight safety.

- Spoofing
  Attacks such as virtual aircraft hijacking virtual trajectory modification or GPS spoofing take advantage of modified messages during the transmission at the physical layer. Methods like overshadowing or bit-flipping are used for this purpose. In both cases, any amount of data of malicious data can be injected without it being notice.

An innovative and economical approach to achieve authenticity and proper integrity to prevent spoofing, deleting and injecting as well as confidentiality to prevent eavesdropping, without re- inventing the implemented protocols, would be the use of a tunneling protocol over existing data link communications systems. This approach should not replace existing data links, like a new universal protocol, but rather complement them in order to achieve a proper level of security.

Difficulty level	**Threat**	**Layer**	**Targer**	**Threat taxonomy classification**
1	Spoofing	Physical and application layer	Ground infrastructure and aircraft	High level
2	Deleting	Physical layer	Ground infrastructure and aircraft	High level
3	Injecting	Physical and application layer	Ground infrastructure and aircraft	Medium to High level
4	Jamming	Physical layer	Ground infrastructure and aircraft	Low to medium level
5	Eavesdropping	Physical and application layer	Aircraft	Low level

TABLE I: Attacking model

III. Cyber Security Issues

- Cyber security workforce
  National Initiative for Cybersecurity Education (NICE) Came up with the Cybersecurity Workforce Framework. This framework insists on recognition by the process of training. Also, accomplishes secure cyber infrastructure and the framework has not included the factor new technologies are rapidly emerging that enhances the challenges in cyber security threats.

- Cyber safety for protecting personal information in computer
  Cyber-safety is a concept that has been used to explain a set of measures, practices, and actions that help in the protection of computer and privacy from various attacks.All devices connected to any company electronic communications network must meet certain security standards. The cyber safety threats can be caused due to viruses, hackers, spyware etc. The virus infects the computer through the

email attachment and file sharing. One infected computer can cause problems to all the computer networks.

- Studies of email virus filtering
  Several studies have been conducted on the filtering of email virus ,Prior study had addressed various existing spam detection methods and finding the useful, precise, and dependable spam detection process. To perform this, the spammer would evaluate the current anti-spam methods. From various studies, it is understood that we will not be able to stop the spam and will be a limit them effectively using Bayesian method when compared to other methods.
- Studies of firewall services
  The performance of personal firewall systems by organizing an arranged walk-through to determine the design factors that could violate the usage standards. The study results indicated that Personal firewalls encounter poor usability that could lead to vulnerabilities in security. This issues has been evaluated by placing the firewalls in the topology of networking design and how to frame the routing tables in the process so that a maximized firewall rule set could be minimal that helps to avoid performance bottleneck and limits the security loopholes.
- Studies of vulnerability scanning
  Intrusion Detection System (IDS) methods is used to identify an attack of a computer network. In order to prevent vulnerable virtual machines network, intrusion detection system is proposed. There is two types of intrusion detection system host based and network based. The evaluation process is distinguished into three sections namely vulnerability adjacency matrix formation, attack processes modeling, and physical consequences analysis.
- Prevention of Cybersafety
  There are seven significant cyber-safety actions which are Running Anti-virus Software, Installing OS/Software Updates, Preventing Identity Theft, Switch on the Personal Firewalls, Prevent Adware/Spyware, protection of Passwords and Backing up Important Files.

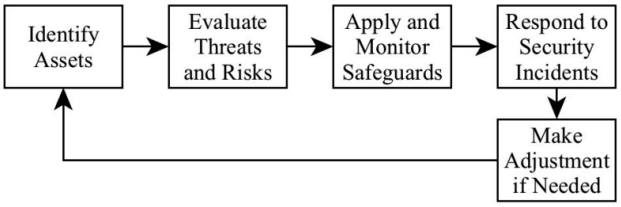


Fig. 1: General operation flow of cybersafety prevention

1) Install OS/Software updates:
   Installing software updates are also known as patches that helps to fix issues of operating system. Many of the latest operating systems are arranged to download updates automatically by default. Once the updates have been downloaded, a confirmation prompt is displayed for installation. Click yes.
2) Running Anti-Virus Software:
   In order prevent computer virus issues install and then run the anti-virus software such as Sophos and check the last updated date.
3) Preventing Identity Theft:
   Don't give out financial account numbers, Social Security numbers, driver's license numbers or other personal identity information unless you know exactly who's receiving it. Protect others people's information as you would your own.
4) Switching on Personal Firewalls:
   Mac OSC and Microsoft Vista have installed built-in firewalls. After turning on the firewall, check it for any open ports which would allow hackers and viruses. Firewalls work as the protection layers between the internet and computers
5) Protecting passwords:
   Make sure that not to share your passwords, and make sure to create new passwords which are hard to guess. Avoid any dictionary works and establish a password by with mixed number, alphabets, and punctuation marks.

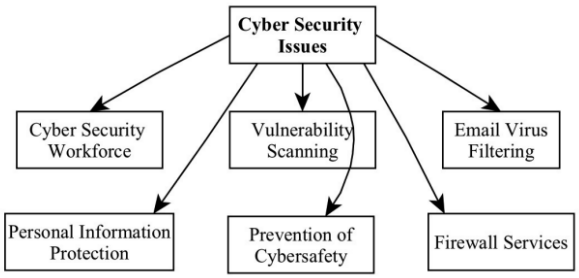


Fig. 2: Cyber security issues

IV. Information Security Awareness

Education programs on information security can play a key role in increasing the overall level of awareness of computer users.The objective is to help organizations prevent being victims of cyber-attacks through social engineering.Social engineering attacks can be performed in several ways including face-to-face, as well as through information technology.Social engineering attacks take advantage of their targets following a typical cycle to gather information or access information systems. The attack cycle typically begins with information gathering from public sources including social media posts, phone books, jobs portals, web pages, among others to develop a mutual relationship with the target. Once a relationship is established, the attacker gets the target to reveal sensitive information such as credit card numbers, login credentials and other valuable data which lead to major cyber attack. To prevent opportunistic attacks, scholars suggest information security awareness and implementing training programs for

employees and members of organizations so that they know how to safeguard their own information and systems and reduce cyber attacks.Many conducted studies confirm that most computer users have a lack of information security knowledge due to insufficient awareness. The Anti-Phishing Work Group (APWG) is one example of a non profit organization running to provide anti-phishing training to increase awareness of information security.

Information security awareness programs help organizations to create a safer computing and data environment in order to keep information properties protected from such malicious attacks. The training of staff enables them to effectively adhere to sound cyber security practises. In terms of the individual awareness level on social engineering threats, several studies have highlighted users' lack of understanding of security and privacy threats associated with personal smart devices they tend to use. As a result, researchers argue that new security and confidentiality actions should be developed with an emphasis on improving the user overall threat awareness caused by similar smart devices. Trust is among the leading security elements that is associated with human's nature. Researchers argue that trust is vital in every aspect of an information security system.

The next generation of learning technologies for information security education, the following methods have the potential for development:

- Online delivery methods including e-mail broadcasting, social media advertising, blogging, and animation.
- Game-based delivery methods, which can challenge, motivate and engage members of an organization.
- Video-based and self-paced delivery methods that allow personnel to study independently.

V. Major Security Goals

There are five major security goals for network security. They are:

1) Confidentiality
   The data or information of any association must be overseen in a way that isn't actually available to unapproved clients. The mystery memory of the substance ofcorrespondence must be there, which assumes an imperative job in security.
2) Availability
   The classified data or information of an association or government workplaces must be put away in a mystery and secure way, however should be obvious to approved clients and must not be available to unapproved clients. Some real clients should likewise be limited.
3) Authentication
   The identity of approved clients must be confirmed when it is important to get to the data or information before getting to the information. There are three manners by which system can confirm the character of the real client. They are passwords, documents and biometric information. Utilizing these confirmation techniques, it is anything but difficult to isolate approved and unapproved clients.
4) Integrity
   The data or information ought not be adjusted or changed amid transmission. The data needs to achieve the goal definitely as it has been sent from the source.
5) Non-repudiation
   The data or information is sent or got; it must be guaranteed that both know about the postponement in sending and accepting information or data. Notwithstanding the principle security goals, there are other optional targets that are important to look after security. They are access and accessibility.

VI. Cyber Threat Intelligence

Threat intelligence is referred to as the task of gathering evidence-based knowledge, including context, mechanisms, in dicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard.Analyzing and sharing threat data and threat information in an effective way requires common representation, standard formats and protocols for sharing, and a common understanding of the relevant concepts and terminology. A solution approach to this need is the use of artificial intelligence (AI) and particularly the use of ontologies.

A. Methodology

1) The Detection Maturity Level Model – DML
   Ryan Stillions proposed the DML model. The model was originally used to describe the maturity of an organization in terms of their ability to consume and act upon given threat information. Threat information can include indicators of compromise, tactics techniques and procedures of an actor (TTPs), threat intelligence reports and many more. The DML model emphasizes the increasing level of abstraction in the detection of cyber attacks, where it is assumed that a security incident response team of low maturity and low skills would be able to detect attacks in terms of low level technical observations in a network, without necessarily understanding the significance of these observations.

2) The Cyber Threat Intelligence Model – CTI
   The Cyber Threat Intelligence model is not hierarchical like the DML model, but mainly a way to represent what types of information are needed for advanced threat intelligence and potential attack attribution. Cyber Threat Intelligence model in the security operations of an organization strengthens the security posture of the organization itself by enabling advanced detective and preventive capabilities.

B. Ontologies

An ontology is a form of knowledge representation that can integrate information coming from different sources. Analyzing the study of several authors we can concludes that

there is not any existing ontology readily available for use within cyber threat intelligence. The main shortcoming is the lack of expressiveness resulting from their poor development and the fact that none of them covers all the relevant data and information (abstraction layers) needed for effective cyber threat intelligence.

Working towards an ontology for cyber threat intelligence is not an easy task. Our research reports the following as the largest difficulties:

- Vaguely defined terminology leads to confusion among experts and additional work to extend or unify ontologies.
- Lack of formal standardized representation of relevant information results in strings of English prose, with no standard pattern. Standardizing well defined taxonomies can eliminate this barrier.
- Lack of coherent relationships between the different layers of abstraction in ontologies. Modular ontologies containing several sub-ontologies need sound relationships between the different data points to leverage the power of semantics and reasoning.

There are several tasks that need addressing in order to create a multi layered cyber threat intelligence ontology. First, formal terminology (definitions) and vocabularies should be described. Second, all the abstraction layers of the cyber threat intelligence model should be included and expressed properly in the ontology. Third, knowledge coming from domain expertise in a structured way should be gathered and formally represented in the ontology to facilitate advanced reasoning based on relationships between data. Fourth, constraints should be defined and constructs should be used in the ontology enabling the reasoning capabilities lying within the OWL language. Finally, the use of subjective logic to model trust in sources and confidence in information.

C. Improving Cyber Supply Chain Security

Cyber supply chain (CSC) is a network of various organizational components connected together with other vendors, suppliers, and consumers for business processes, goals, and objectives.Cyber threat intelligence(CTI) provides an organization the capability to identify,gather, analyze threats and the associated risks so that CSC organization can forecast the existing and future threat trends and manage the cybersecurity risk.Factors such as known-known attacks, Known Unknown attacks and Unknown-Unknown attacks that are impacting on the evolving threat landscapes and affecting the CSC.

1) Integration of CTI for CSC
   Integrating CTI in CSC systems has become integral due to factors such as an increase in the use of a sophisticated cyber physical system (CPS).It considers CTI core cyber threats concepts such as indicators, TTPs, incident, and threat actors and links them with CSC concepts.CTI is relevant for CSC as it provides detailed knowledge and understanding of the threat actor's motives. Thus, without the required integration of CTI lifecycle approach into CSC development, it will be highly unlikely to effectively mitigate the attacks, risks, vulnerabilities, and exploits that are currently prevailing on CSC systems.

2) Control Mechanisms
   Controls are a security mechanism put in place to mitigate specific cyber threats.It includes
   - Create Inventory and Control of Hardware Assets.
   - Create Inventory and Control of Software Assets.
   - Continuous Vulnerability Management.
   - Controlled Use of Administrative Privileges.
   - Secure Configuration for Hardware and Software on Mobile Devices,Laptops,Workstations and Servers.
   - Maintenance,Monitoring and Analysis of Audit Logs.

   The CTI gathered will inform the strategic, tactical and operational management roles and responsibilities. Operational level managers could use the CTI indicators, to determine details of attacks and TTPs, relate attacks and provide remediation from the control statement.

VII. Conclusion

Today, organizations are greatly dependent on information systems. This reliance has led to being vulnerable to information security threats that put systems at risk. Furthermore, social engineering fraud has been rising significantly with advancements in technology. Criminals are getting more sophisticated in finding new ways to attack. As a result, organizations have been increasing their investments in cyber security initiatives to safeguard their data. Cyber Threat intelligence gathering is a proactive way to assist organizations to gather, analyze and disseminate intelligence, CTI could assist in determining whether an attack could occur in the near future.From the review, it was found that majority of the studies have been conducted on the email security, firewalls, and vulnerabilities. Yet, not many studies from the perspective of password security. Therefore, there is a need for more studies in terms of technics and models from this perspective to ensure that passwords are protected.There by we can ensure cyber security among individuals and organizations.

References

[1] Kutub Thakur, Meikang Qiu, Keke Gai, Md Liakat Ali . An Investigation on Cyber Security Threats and Security Models . 2015 IEEE 2nd International Conference.
[2] Richa Adlakha, Shobhit Sharma, Aman Rawat, Kamlesh Sharma Department of Computer Science & Engineering Cyber Security Goal's, Issue's, Categorization & Data Breaches .2019
[3] Philipp Ortner,Holger Flühr,Erich Leitgeb Cyber Security and Information Exchange Analysis of Automatic Dependent Surveillance Broadcast.
[4] Siri Bromander,Vasileios Mavroeidis Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence,2017
[5] Geoffrey Skinner,Hussain Aldawood Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review,2018
[6] Ezer Yeboah-Boateng ,Abel Yeboah-Ofori, Shareeful Islam Cyber Threat Intelligence for Improving Cyber Supply Chain Security,2019
[7] SANS, "Security Intelligence: Attacking the Cyber Kill Chain," https://digital-forensics.sans.org/blog/2009/10/14/security-intelligence attacking-the-kill-chain/.

[8] T. Casey , "Threat Agent Library Helps Identify Information Security Risks," Intel White Paper, September, 2007.

[9] J.–W. H. Bullée, L. Montoya, W. Pieters, M. Junger, and P. H. Hartel, "The persuasion and security awareness experiment: reducing the success of social engineering attacks," J. Experimental Criminology, vol. 11, no. 1, pp. 97–115, Jan. 2015.

[10] W. Kearney and H. Kruger, "Considering the influence of human trust in practical social engineering exercises," in Proc. Information Security for South Africa (ISSA '14), Johannesburg, South Africa, 2014, pp. 1–6.

[11] J, Freidman, and M Bouchard. "Cyber Threat Intelligence Guide: Using Knowledge About Adversary to Win The War Against Targeted Attacks." iSightPartners. 2018.

[12] Z, Porkorny, "What Are the Phases of The Threat Intelligence Lifecycle?" The Threat Intelligence Handbook. A Practical Guide for Security Teams to Unlock the Power of Intelligence. 2018.

[13] S. Ahmed, M. Elsholkami, A. Elkamel, J. Du, E. Ydstie, and P. Dou glas. Financial risk management for new technology integration in energy planning under uncertainty. Applied Energy, 128:75–81, 2014.

[14] Y. Badr, F. Biennier, and S. Tata. The integration of corporate security strategies in collaborative business processes. IEEE Trans. on Services Computing, 4(3):243–254, 2011.

[15] RTCA, Inc., "RTCA DO-242 Minimum Aviation System Performance Standards for Automatic Dependent Surveillance Broadcast (ADS B)," 13 December 2006.

A Literature Review of Mobile Cloud Computing

Anupama P N, Reshma Reghu and Steny Baby
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Siji K B
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: siji.k.b@vidyaacademy.ac.in)

***Abstract*—Mobile cloud computing (MCC) which combines mobile computing and cloud computing, has become one of the industry buzz words. As MCC is still at the early stage of development, it is necessary to grasp a thoroughunderstanding of the technology in order to point out the direction of future research. With the latter aim, this paper presents a review on the background and applications of MCC and future research trends. A brief account on the background of MCC: from mobile computing to cloud computing is presented and then followed with a discussion on system architecture and programming models. It then analyses the basic MCC service models and access schemes for MCC. The rest of the paper analyses the taxonomy, challenges and opportunities of Mobile Cloud Computing, and points out promising future research directions.**

***Index Terms*—Keywords: Mobile Cloud Computing, Mobile Computing, Cloud Computing, IaaS, PaaS, SaaS**

I. Introduction

In recent years, applications targeted at mobile devices have started becoming abundant with applications in various categories such as entertainment, health, games, business, social networking, travel and news. The reason for this is that mobile computing is able to provide a tool to the user when and where it is needed irrespective of user movement, hence supporting location independence. Indeed, 'mobility' is one of the characteristics of a pervasive computing environment where the user is able to continue his/her work seamlessly regardless of his/her movement. However, with mobility comes its inherent problems such as resource scarceness, finite energy and low connectivity. 'The future belongs to services that respond in real time to information provided either by their users or by nonhuman sensors'. Real time applications are just one type of mobile applications that demand high levels of responsiveness, that in turn, demand intensive computing resources. Some mobile applications, such as location based social networking, process and make use of the phone's various sensor data. However, extensive use of sensors, such as obtaining a GPS reading, is expensive in terms of energy and this limits the mobile phone in providing the user a better service through its embedded sensors. Furthermore, consider applications that require extensive processing– image processing for video games, speech synthesis, natural language processing, augmented reality, wearable computing— all these demand high computational capacities thus restricting the developers in implementing applications for mobile phones.

With the explosive growth in mobile applications, platforms and end user demands, limitations at the mobile device end (e.g., computation and storage capacity, energy, shared wireless medium) significantly impede further improvements in application quality of service (QoS)- typical measures of QoS include user experienced delay, service reliability/availability and information privacy. Mobile cloud computing has recently attracted significant attention from both academia and industry. Mobility is now an extremely popular phrase and quickly growing aspect in today's computing area. Mobile computing is referred to as a sort of human-computer conversation with which a computer are estimated to be delivered during normal usage. Mobile computing draws on an accumulation of a few significant concepts: hardware, software and communication. Cloud Computing is comparable to distributed computing. In the latter, the resources are distributed over the network, the computing in the information is completed in things and the outcomes are built-in to obtain the end results 4 .By doing this we decrease the issue of huge computations being done only one site. The idea of distributed computing work is used on cloud computing. There is absolutely no sole consensual meaning for cloud computing because various developers and institutions describe it in numerous ways6 .The primary a part of the cloud computing system is this cluster system, called the Cloud13 .Cloud organizes the numerous resources of the computers based on a particular management plan and offers safe, reliable, quick, suitable and transparent services for example information storage, retrieval and computing to clients. This liberates the client of the mentioned duties .

The phrase "Mobile Cloud Computing" was introduced shortly after the idea of integration of mobile and cloud computing released few years back. It is often drawing the attentions of entrepreneurs as a prosperous business option that decreases the growth and running price of mobile applications, of mobile as a brand new technology to attain rich experience of a number of mobile services at low priced, and also of experts as a ensuring solution for environmentally friendly IT2 . This part provides an overview of MCC including definition and architecture of MCC.

II. MOBILE CLOUD COMPUTING

There are several existing definitions of mobile cloud computing, and different research alludes to different concepts of the 'mobile cloud':

1) 1. Commonly, the term mobile cloud computing means to run an application such as Google's Gmail for Mobileon a remote resource rich server (in this case, Google servers) while the mobile device acts like a thin client connecting over to the remote server through 3G. Some other examples of this type are Facebook's location aware services, Twitter for mobile, mobile weather widgets etc.
2) 2. Another approach is to consider other mobile devices themselves too as resource providers of the cloud making up a mobile peer-to-peer network. Thus, the collective resources of the various mobile devices in the local vicinity, and other stationary devices too if available will be utilized.

A. APPLICATIONS

Some of the key classes of mobile application include: Mobilelearning could be described as, using mobile technology to improve or augment people knowledge, skills and behaviour through education, training, or performance. It adds a new level of portability to the learning process through the use of tablets and smart devices. Mobile devices are being used extensively in business and commerce for a wide range of purposes. The obvious traditional Mobile commerce applications include mobile transactions and payments, mobile tickets and advertisements through messaging services(SMS or MMS messages).Mobile health and wellness applications benefit from the integration of more sophisticated body area network systems that provide improved real time sensing capability of health parameters. Mobile Cloud for Assisting Healthcare (MoCAsH) is a prototype implementation of m-healthcare information management system based on cloud computing for mobile users. Mobile social medias, many of the MCC applications are now assuming a presence in the social media world where they can be easily discovered and deployed we use the term Mobile Social Network (MSN) to broadly refer to such applications . Today, mobile applications are often developed based on social networks such as Facebook. In fact these social networks serve as defacto means of entry login into many mobile services which bridge the gap between virtual social communities and physical world information and services. By using LBSs mobile users could get localized information like maps, driving directions, search of point of interests, sharing data etc.

III. MCC SYSTEM ARCHITECTURE

Current cloud computing providers typically allow customers to rent computation and storagesuch that customers can start instances of their cloud applications as VMs within the provider's cloud ofservers. Cloud providers may provide additional services, such as backup and traffic accounting, to ease the process of managing VM instances. The distribution of the VM instances is largely transparent to the customers, and cloud providers mainly focus on providing guarantees of CPU time, memory usage, storage, server availability, networking throughput, etc. However, some cloud providers offer customers the additional ability to choose geographically from among a small number of data centers where their VM instances will run. Today's mobile applications have already begun to adapt to cloud computing. A common theme emerging from the large wave of mobile applications developed for smart phones such as the iPhone and Android is that these mobile applications are often linked to server instances operating in the cloud.

A. CLOUD SERVICE MODELS

The term "cloud services" refers to a wide range of services delivered on demand to companies and customers over the internet. These services are designed to provide easy, affordable access to applications and resources, without the need for internal infrastructure or hardware. There are the following three types of cloud service models –

1) Infrastructure as a service (IaaS)
2) Platform as a service (PaaS)
3) Software as a service (SaaS)

1) Infrastructure as a service (IaaS): IaaS is also known as Hardware as a Service (Haas). It is a computing infrastructure managed over the internet. The main advantage of using IaaS is that it helps users to avoid the cost and complexity of purchasing and managing the physical servers.
Characteristics of Iaas:

1) Resources are available as a service
2) Services are highly scalable
3) Dynamic and flexible
4) GUI and API-based access
5) Automated administrative tasks

Examples: DigitalOcean, Linode, Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE), Rackspace, and Cisco Metacloud.

2) Platform as a service (PaaS): PaaS cloud computing platform is created for the programmer to develop, test, run, and manage the applications.
Characteristics of PaaS:

1) Accessible to various users via the same development application.
2) Integrates with web services and databases.
3) Builds on virtualization technology, so resources can easily be scaled up or down as per the organization's need.
4) Support multiple languages and frame work.

Example: AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, Magento Commerce Cloud, and OpenShift.

3) Software as a service (SaaS): SaaS is also known as "on-demand software". It is a software in which the applications are hosted by a cloud service provider. Users can access these applications with the help of internet connection and web

browser.
Characteristics of SaaS:

1) Managed from a central location
2) Hosted on a remote server
3) Accessible over the internet

B. PROGRAMMING MODELS

There are several factors to consider for developing programming models. First, we need to find out the objective. Various objectives are offload computation, reduce latency, minimize energy consumptions, etc. Second, we need a profiler to understand the resource usage of the components, andthe impact of offloading. Third, we need a solver to decide what to offload.

The first is the recently proposed Clone Cloud. It is proposed that the mobile device will have a clone copy in the cloud. The two work in synergy to enhance the application experience while minimizing resource consumption. This programming model allows a mobile client to fully utilize the cloud resources. The optimization solver decides on what executions should be offloaded based on a dynamic profiler and a static analyzer.The second one makes use of RPC to remotely execute resource intensive methods. MAUI proposes such a programming model. Since many media applications make use of standard components for face recognition, gesture recognition, object and pose recognition, packaging them as cloud services with standard APIs can be more appealing. This motivates our REST ful programming model. For REST ful model, there is no state kept in the cloud. Whenever a computing task is needed, the mobile device just invoke a function with appropriate parameters through http or https protocol.

IV. BASIC MOBILE CLOUD COMPUTING SERVICES

Cloud computing providers will provide a set of basic services for mobile computing. There arethree types of services. The first one is what we refer as platform services, the second is application services, and the third is context-rich support services.

A. Platform services

Platform services include computing, storage, database, memcache, content distribution. Out of the basic platform service, one can already build very useful applications such as file backup service, and filesyncing service, etc.

B. Application services

Public cloud provider can also offer a set of essential application services. For example, people may not trust each individual applications and thus, may not reveal their location information. This canhamper the development of location based services. If mobile devices are using the cloud services, thenthere is prior trusted relationship. So it is easier to trust the cloud provider for location privacy. Thus, apresence service can be an essential service so that any application that needs location information cantalk to the presence service. So it is easier to trust the cloud provider for location privacy. Thus, apresence service can be an essential service so that any application that needs location information cantalk to the presence service. The presence service will implement location privacy policies according towhat are stipulated by the mobile subscribers.

C. Context-rich services

Many mobile applications will become more personalized, and more context aware, recognizing not only the location of the user and the time of day, but also a user's identity and their personal preferences. To support these mCloud services, we believe mCloud providers need to provide a set of context-rich support services. Application developers can use these context-rich support services as building blocks to build a large class of new mCloud services. Based on these contextual clues, a layer of cloud recommendation services can be built that creates output that is tailored to an individual or set of individuals with those contextual characteristics. For example, some applications have begun to combine together mobile location with social networks to generate multimedia content, e.g. a song playlist or a recommended video, that is tailored to the individual or individuals who are nearby an audio jukebox or video screen that is aware of their presence.

V. ACCESS SCHEMES FOR MOBILE CLOUD COMPUTING

The concept of Mobile Cloud Computing (MCC) intends to make the advantages of Cloud Computing available for mobile users but will provide additional functionality to the ”cloud” as well. Mobile Cloud Computing (MCC) will help to overcome limitations of mobile devices in particular of the processing power and data storage. A number of new business roles and technical functionalities might be provided by mobile clouds.Today, context and location information is already used by a broad variety of applications, in particular context-aware services for mobile terminals. These services exploit data collected from terminal sensors (e.g., GPS, gyro, proximity detectors), sensor networks, or network sensors measuring network status and load. Classical so-called ”intelligent access” concepts are assuming that all categories of dynamic context information (user profiles, terminal status and sensor information, external sensor networks, and network status) can be used to optimize the mobile access. The requirements on the wireless connectivity for Mobile Cloud Computing differ in a number of important details from these classical heterogeneous access scenarios:

- MCC requires an “always-on” connectivity for a low data rate cloud control signaling channel.
- MCC requires an “on-demand” available wireless connectivity with a scalable link bandwidth.
- MCC requires a network selection and use that takes energy-efficiency and costs into account.

The solution is that we deploy MCC in a heterogeneous access scenario with a wide range of different radio access technologies available, e.g., GPRS, WCDMA/HSPA, LTE,

WiMAX, cdma2000, WLAN, and that context information, such as terminal locations and capabilities and user profiles, can be used by the mobile cloud controller to locally optimize the access management.

In a more and more fragmented and heterogeneous wireless network landscape and with an increasing demand for high data rates or even on-demand scalable transmission bandwidth as for MCC scenarios, the need for efficient network access management across different radio access technologies (RATs), referred to as "heterogeneous access management" (HAM), becomes eminent. Radio resources are of course limited due to physical, technology-specific and regulatory constraints. Context, either provided by terminals, network nodes or from sensors deployed in the user's environment, will be utilized to significantly reduce the waste of scarce radio resources and efficiently manage wireless access across heterogeneous RATs.The issue of network selection in an environment of diverse wireless access possibilities is widely addressed in today's research. HAM describes a concept for controlling radio resource allocation and utilization, as e.g. bandwidth, power, etc., across various RATs, where the main goals are to optimize overall system performance and to enable seamless mobility. HAM decisions and actions may be influenced by network operator policies, service level agreements ,user preferences, user location information or as a result of sophisticated resource utilization analysis.

The heterogeneity of today's wireless access possibilities-imposes challenges for efficient access and resource managementacross different RATs. The proposed IRNA concept, accounts for the characteristics and statusof each RAT while simultaneously trying to accommodateusers, who face different environmental conditions, with thebest possible end-to-end performance. The Context Management Architecture(CMA) is designed to acquire, manage, and distribute context information and to control the context quality required for IRNA purposes. This functionality requires the following core architectural components:

- Context Provider
- Context Broker
- Context Consumer.

Moreover, a context quality enabler (CQE) is responsible forcontrolling quality of context. The Mobile Cloud Controllercan but does not need to operate the Context Broker(platform).We assume that the Mobile Cloud, on the one hand, providescontext aware services and, on the other hand, uses contextinformation for the intelligent radio access management.

The Context Provider(CP) is the logical point in the CMAwhere context information originates from and is provided toother entities of the architecture. The types of context dataa CP provides can be classified in (relatively) static data and dynamic data. Static data forinstance includes terminal capabilities, user preferences or userinformation retrieved from social communities on the web. However, for the purpose of managing IRNA, dynamic contextinformation, such as user location and movement (i.e. speedand direction) and network conditions have a higher relevance.

The Context Broker (CB) takes on the role of a middleman. Its main functionality is to maintain a registry of available CPsand their capabilities based on the announcement sent by CPs. Based on that registry, it can provide a CP look-up service toentities searching for certain context data. In addition, the CBcan provide context data itself by forwarding the data it hasreceived from CPs. Two different communication modes areavailable for requesting entities. In the asynchronous mode,context is forwarded if a specified condition or event comestrue. In the synchronous mode, arequest for context information is instantly answered by theCB.

The Context Consumers (CC) are entities that use context data asan input for their actual functionality. Network services, applicationsfor end users, service enablers or actuators in wirelesssensor and actuator networks are exemplary consumers of context information. The CCs we focus on here consist ofdifferent components of the IRNA framework . CCs can request context from the CB or they can invoke anURI of a CP at the broker. Consequently, CCs can establish adirect communication with the according CP.

Context-aware systems managing IRNA in a dynamic, multi-technology, multi-vendor environment have to performa multitude of tasks to optimize network utilization. In orderto make the right decisions in terms of network selection,link adaptation, prediction of user behaviour, and multicasting, IRNA modules require context data of high quality. The integration of a Context Quality Enabler (CQE) into the CMAassures that only context information that adheres to thestated quality criteria will be used for IRNA management. The conceptual outline of the architecture allows for controlling context quality according to the needs of the CCs. Depending on the actual implementation of the CQE, context quality can be controlled with respect to different criteria.

In order to evaluate the performance of radio access schemesfor Mobile Cloud Computing, a radio network simulator takingcontext information into account is required. The proposed Context-Aware Radio Network Simulator (CORAS) incorporates static as well as dynamic context models to simulate the influence of context quality on HAM and users' end-to-end performance. Context quality is controlled with respect to the criteria availability, accuracy, delay, confidence, and relevance forassessing link quality. For instance, current network capacity is modelled based on past capacity and instantaneous capacityestimation. However, data on instantaneous capacity is onlyavailable with a certain delay. The smaller this delay, thebetter the quality of context information and, hence, the better the instantaneous resource utilization. Further, preliminary simulation results have shown that adelayed user movement prediction with respect to speed anddirection increases the number of handovers. Thus, worse context quality results in less effective resource utilization. As a conclusion we can say that by removing the limitations of mobiledevices with respect to storage and computing capabilities andproviding a new level of security by a centralized maintenanceof security-critical software it is expected that it will find broad acceptance on the businessas well as consumer side.

VI. A TAXONOMY OF MOBILE CLOUD COMPUTING

Our criteria for defining the taxonomy is based on the key issues in mobile cloud computing, and how they have been tackled in academia. We focus on:

1) Operational level issues
2) End user level issues
3) Service and application level issues
4) Privacy, security and trust
5) Context-awareness
6) Data management as the main areas.

A. Operational issues

Operational issues refer to underlying technological matters such as the method of offloading computations, cost–benefit models that aid in taking the decision to offload or not, how the mobility of devices is managed/supported, and connection protocols used.

1) Method of offloading : The main operation in any mobile cloud would be the offloading of jobs that take place from the resource constrained mobile device to the cloud. Because of issues such as the physical distance separating the mobile device and the cloud and the heterogeneity of the underlying systems, different research has tackled this in a variety of ways. Current research discusses offloading methods in three main directions; Client–ServerCommunication methods, Virtualization, and Mobile agents. In the Client–Server Communication process communication is done across the mobile device(offloader) and surrogate device via protocols such as Remote Procedure Calls (RPC), Remote Method Invocation (RMI) and Sockets. Both RPC and RMI have well supported APIs and are considered stable by developers. However, offloading through these two methods mean that services need to have been pre-installed in the participating devices. This is a disadvantage when considering the ad hoc and mobile nature of a mobile cloud and restricts the mobility of users if in the vicinity of devices that do not support the needed services. Virtual machine (VM) migrationrefers to transferring the memory image of a VM from a source server to the destination server without stopping its execution. In such a live migration, the memory pages of the VM are pre-copied without interrupting the OS or any of its applications, thereby providing an illusion of seamless migration. This method ensures that no code changes are needed when programs are offloaded, and provides a relatively secure execution, since the VM boundary insulates the surrogate device. However, VM migration is somewhat time consuming and the workload could prove to be heavy for mobile devices. Scavenger is another framework that employs cyber-foraging using WiFi for connectivity, and uses a mobile code approach to partition and distribute jobs. It also introduces a scheduler for cost assessment. Its method of cost assessment is based on the speed of the surrogate server and it uses a benchmarking method to do this. Using its framework, it is possible for a mobile device to offload to one or more surrogates and its tests show that running the application on multiple surrogates in parallel is more efficient in terms of performance. However, it does not discuss fault tolerance mechanisms and since its method is strictly about offloading on surrogates and not sharing, it is not really dynamic. Also its surrogates are all desktops and it is unclear if Scavenger is too heavy to run on mobile phones.

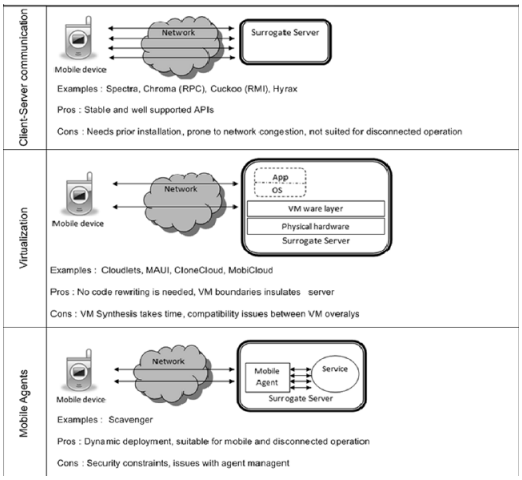

Client-Server communication
Network
Surrogate Server
Mobile device
Examples : Spectra, Chroma (RPC), Cuckoo (RMI), Hyrax
Pros : Stable and well supported APIs
Cons : Needs prior installation, prone to network congestion, not suited for disconnected operation
Virtualization
Network
App
OS
VM ware layer
Physical hardware
Surrogate Server
Mobile device
Examples : Cloudlets, MAUI, CloneCloud, MobiCloud
Pros : No code rewriting is needed, VM boundaries insulates server
Cons : VM Synthesis takes time, compatibility issues between VM overalys
Mobile Agents
Network
Mobile Agent
Service
Surrogate Server
Mobile device
Examples : Scavenger
Pros : Dynamic deployment, suitable for mobile and disconnected operation
Cons : Security constraints, issues with agent managent

Fig. 1. Overview of Offloading Methods used in Mobile Cloud Computing Systems

2) Cost–benefit analysis: It is important to analyze the costs of offloading on to the cloud such as time, energy and monetary, versus monolithic execution/storage beforehand.

3) Mobility management : One of the key issues encountered in a mobile cloud is the design of intelligent mobility management techniques that support user mobility while providing a seamless service promoting task redundancy. Mobility is one reason in mobile clouds for disconnectivity, and disconnection. Although research has been done in location management in wireless networks, here, we focus on the methods followed to manage and support mobility in mobile cloud computing systems. Determining a device's current location is helpful to assess its potential to move away from or towards the active mobile cloud. Work on localization primarily falls into either infrastructure based techniques, or peer-based techniques. Infrastructure based methods use technologies such as GSM, Wi-Fi, ultra-sound with RF, GPS, RFID, and IR. Among these, current GPS very rarely works indoors, and the more accurate techniques need additional infrastructure and need dense deployment of access points as well.

4) Connection protocols: The current mobile cloud computing research uses a variety of connection protocols for communication including Wi-Fi, Bluetooth, and 3G, though the majority has employed Wi-Fi for many reasons. Wi-Fi was initially intended as replacement for cabling for resource and peripheral sharing (such as printers, shared storage devices) among PCs, terminals etc. for wireless local area networks (WLANs).Wi-Fi has a longer range, with a radius within 100m and supporting up to 11 Mbps data rates. Bluetooth on

the other hand, was intended for non-resident equipment and applications such as wireless headsets etc. wireless personal area network (WPAN), and is characterized by its low power requirements and low-cost transceiver chips. The range for Bluetooth is typically in a radius of 10 m, depending on the device class, power, and physical obstacles in the environment. 3G (third generation mobile telecommunications) is a technology for mobile service providers and it shares the basic business model with that of the telecommunications services model. The infrastructure is owned and managed by the service provider and sold to customers typically on a monthly usage basis. Although the focus of cellular technology has been voice telephony, data services has also started to attract attention.

B. End User issues

End users issues relate to issues that directly involve users such as incentives for participating, interoperability and cost. When using a mobile cloud, one of the key challenges experienced by the end users is the transaction infrastructure.

C. Service and application level issues

Service and application level issues relate to the factors concerned with performance measurements of the system, and the QoS of the system. Fault-tolerance is a highly important aspect in a mobile cloud, even more so than a conventionalcloud because of the mobile nature of the devices. Disconnection can happen due to user mobility asdevices enter and leave a network. Running out of battery power, network signal loss, or hardware failuresare other common factors. A majority of Application Programming Interfaces used to build mobile applications targeted forcloud computing are based on service oriented architecture such as REST and/or SOAP. Mobileapplications are able to connect to and request services hosted on a remote cloud through interfaces. However, mobile Web services need to consider additional constraints other than standard Web services: frequent loss of connectivity, low computational resources, and low bandwidth.

D. Privacy, security and trust

Whether offloading intensive computations, or data storage, using the cloud for mobile devices does pose questions of security and trust issues. Because of the low capacity of mobile device storage, many users are starting to store data such as contacts, calendars and SMS on clouds. However, these cloud services are stated to be vulnerable and users may lose their data if the services go out of business, or simply if the services fail due to technological problems.Mobile cloud computing inherits the security threats of conventional cloud computing in cases when the definition of mobile cloud means to connect mobile devices to a remote cloud. In this case, the remote cloud server would be the same as a conventional cloud computing provider, making the general cloud security threats valid. At the same time, mobile clouds present a group of issues that are particular to mobile devices offloading jobs through wireless communication channels. Furthermore, security concerns that are specific to mobile devices such as battery exhaustion attacks, mobile botnets and targeted attacks should also be considered.

E. Context-awareness

Systems with context-awareness are able to use contextual information to change and automatically reconfigure their configurations to adapt to the context [44]. This behaviour is very useful in the case ofmobile systems since these deal with an execution environment that is subject to constant change. In thecase of mobile cloud computing, context awareness can be used in forming resource clouds as well asprocessing information. For example, a device can infer its location through GPS, Bluetooth, or some otherforms ofpositioning and use that information to prepare itself for upcoming processing. A Context Manager component is deployed to sense context information and store it to beused for other components such as Application Manager, which launches, intercepts and modifies anapplication according to the current context. Location and number of nearby devices in the vicinity are thekey contexts,with location information used for mobility traces, and number of devices used to aid the forming of themobile cloud. Thus, the system is made aware if a new device enters the resource pool, or leaves it, therebyleading to better scalability and content distribution.

F. Data management

For many cloud users and providers, managing data on the cloud raises many complications. In mobile cloud computing, as the name itself suggests, data that would traditionally be only accessible onlyto the mobile device's owner, would now be stored on, accessible to, shared with external devices or users. For many mobile users, this raises privacy and security questions. Also, data representation in mobiledevices vary, and in a heterogeneous mobile cloud, this would lead to problems with portability andinteroperability. Computations in a mobile cloud would be spread across a distributed file system, wheremultiple devices may needto access and modify files. Furthermore, special considerations must be given to accessing files frommobiles over wireless networks. In a mobile cloud, users geographical locations are not fixed, andbandwidth must be conserved because of data access costs. However, it is also vital that mobile databasescontain policies to safeguard against data loss while ensuring it conforms to mobility constraints. One of the key concerns for people about using a mobile cloud is that their personal data onmobile device could be stored on, or accessed by the cloud. A mobile device contains contact lists, textmessages, personal photos and videos, calendars, location information, and these data can reveal manythings about someone's personal life. However, a personal computer also stores many such personal datasuch as photos and other multimedia, chat logs, emails, passwords, financial records or access to suchrecords, calendar and contact lists. Today, thousands of monetary transactions are being done in onlineshopping sites such as eBay, and Amazon. Therefore, the risks involved in mobile cloud are not necessarilygreater than those involved in traditional clouds. However, the issue here is

whether the means of handlingthose risks have been properly implemented in the mobile cloud. Despite reservations, people do tend touse their mobile devices with the cloud. Some recent examples of mobile cloud storage are Apple's iCloud, Google Drive and Dropbox. In the context of the mobile cloud, mobile databases need to be lightweight, and requirethe ability to download data from a remote repository and execute on this data even in a disconnected state, and also should be able to synchronize the modified data during the downtime with the enterprise wheneverthe network becomes available again. Furthermore, databases for mobile clouds need to have a quick startup time since faults in mobile devices can be more frequent than for a fixed host. Security constraintsregarding access and real time processing are also important.

VII. Challenges

Though it may sound that the use of MCC to develop applications is like a walk in the park, it is not so in practice. A few challenges that crop up while using this technology to develop apps include,

1) Less network bandwidth Carrying out deployment using MCCs requires the communication to be continuous. This means that a developer may face problems if the network being used is wireless. This is because wireless networks tend to be less reliable or possess low bandwidth. For example, 3G, Wi-Fi, or 4G networks. Therefore, the speed of the applications is much slower in comparison to wired networks. While 5G networks remain a ray of hope, it is much too early to decide its effectiveness.
2) Service availability Mobile users may attain a very low-frequency signal, hindering the speed, as well as the storage capacity of the application. Moreover, users also experience issues like breakdown, transportation crowding, and lack of coverage.
3) Hardware Issues Mobile phones, even with the latest technology, have a finite source of energy, i.e., batteries. Cloud-based apps increase the use of the battery and would, therefore, drain it much more quickly. This can hinder MCC development as the user base can potentially decline along with an increase of complaints regarding the impact on the battery life.
4) Operating System Issue The applications created using MCC will function on different operating systems. Therefore, the application must be compatible with operating system platforms like Android, iOS, and Windows Phone. To do so, the development team must possess knowledge regarding an IRNA or Intelligent Radio Network Access technique.
5) Security Issues The management and identification of threats have proved to be a challenging task. This is because MCCs function on a wireless network. Therefore, there are more chances of overlooking or the general absence of network information. Moreover, with multiple hand-offs within the architecture and a general lack of multi-layer security, vulnerabilities are high. The security related issues stem from vulnerabilities in the MCC architecture. With multiple users accessing the clouds there is a threat to the safety of data. Say if the security of one of the user data is breached then there are risks at other users as well.

VIII. Conclusion

Today's mobile applications are demanding compute intensive capabilities such as speech recognition, natural language processing, computer vision and graphics, machine learning, augmented reality, planning and decision making. These demands will not be met solely by making more powerful mobile devices. By thus empowering mobile users, mobile computing will be able to break free of the fundamental constraints that have been keeping us from transform many areas of human activity.

Mobile cloud computing aims to empower the mobile user by providing a seamless and rich functionality, regardless of the resource limitations of mobile devices. Although still in its infancy, mobile cloud computing could become the dominant model for mobile applications in the future. The vision of MCC is an autonomous digital environment in which different mobile devices obtain their computation, storage, services and other resources autonomously and efficiently anytime and anywhere.

References

[1] M. Satyanarayanan, Mobile computing, Computer 26 (1993) 81–82.
[2] L. Siegele, Let it rise: a special report on corporate it, http://www.economist. com/node/12411882, 2008.
[3] W. Vogels, A head in the clouds the power of infrastructure as a service, in: Proceedings of the 1st Workshop on Cloud Computing and Applications, CCA'08.
[4] J. Cheng, R.K. Balan, M. Satyanarayanan, Exploiting rich mobile environments, Technical Report, 2005.
[5] R.E. Frederking, R.D. Brown, The pangloss-lite machine translation system, in: Proceedings of the Second Conference of the Association for Machine Translation in the Americas, pp. 268–272.
[6] E.E. Marinelli, Hyrax: cloud computing on mobile devices using MapReduce, Masters Thesis, Carnegie Mellon University, 2009.
[7] Y. Sasaki, Y. Shibata, A disaster information sharing method by the mobile servers in challenged networks, in: Advanced Information Networking and Applications Workshops, WAINA, 2012 26th International Conference on, pp. 1048–1053.
[8] J. Dean, S. Ghemawat, MapReduce: simplified data processing on large clusters, Communications of the ACM 51 (2008) 107–113.
[9] L. Xinhui, L. Ying, L. Tiancheng, Q. Jie, W. Fengchun, The method and tool of cost analysis for cloud computing, in: Proceedings of IEEE International Conference on Cloud Computing, CLOUD'09, pp. 93–100.
[10] M. Kristensen, Scavenger: transparent development of efficient cyber foraging applications, in: Proceedings of the IEEE International Conference on Pervasive Computing and Communications, PerCom.

Green Computing: The Future of Computing

Anjitha M P, Aleena Vincent and Dijina P D
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Salkala K S
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: salkala@vidyaacademy.ac.in)

***Abstract*—Green computing or green IT, refers to environmentally sustainable computing or IT. Green computing is defined as "the study and practice of designing, manufacturing, using, and disposing of computers, servers, and associated subsystems such as monitors, printers, storage devices, and networking and communications systems efficiently and effectively with minimal or no impact on the environment. The goals of green computing are reduce the use of hazardous materials, maximize energy efficiency during the product's lifetime, and promote the recyclability or biodegradability of defunct products and factory waste. Research continues into key areas such as making the use of computers as energy-efficient as possible, and designing algorithms and systems for efficiency-related computer technologies.**

***Index Terms*—Green Computing, Green IT, Energy Efficiency, E-waste, Environment, Computational Intelligence, Hybride Systems, ICT**

I. Introduction

The research and practice of ecologically sound computing or IT are known as green computing, green Information technology. The objectives of green computing are comparable to those of chemical synthesis: decrease the use of dangerous chemicals, maximize energy efficiency throughout the product's lifecycle. The maximize recyclability or biodegradability of obsolete manufacturing and production garbage. Green computing is critical for all systems, from personal computers to large-scale information centres.

The primary goal of adapting green computing is to facilitate the organizations to use various computer resources proficiently by maintaining and increasing its overall performance. Green computing is the branch of an environmental science which enables the responsible use of computer resources, servers and its associated subsystems like display devices, various input and output devices. The green-Computing, as defined in the Official Journal of the French Republic in 2009, is an eco-friendly information technology and communication technique by using which we can reduce the harmful effects of human activity on the atmosphere.

Green computing can be thought of as the responsible use of computers and their corresponding resources for achieving environmental and ecological goals. More generally, it is also defined as the study of designing, manufacturing/engineering, utilizing and discarding of computing devices in a way that decreases the impact on the environment for the benefit of society. For this reason, it is of upmost importance to consider this area in every aspect of human life. At the same time, the use of computational intelligent techniques has also become very important for everyday life, ranging from domestic appliances to autonomous systems and automated medical diagnosis.

The waste of hardware materials contains chemicals which are toxic. A computer left for so many hours will produce high electronic cost as well as it produces carbon-di-oxide in the atmosphere. But now-a-days, more and more computers producing unnecessary greenhouse gases every year which can cause global warming. Green computing is designing, manufacturing, using and disposing of computers, server and associated subsystems efficiently with minimal impact on the environment.

It starts from design to manufacturing to use of equipment and then safely disposing-off computers, related devices, networking and communications equipment efficiently and effectively with negligible or no impact on the environment. First and most conclusive research on computing shows that Carbon Dioxide (CO2) and other emissions are causing global climate and environmental damage. Preserving our beloved

planet is a main and legitimate goal because it aims to preserve life. It has given extreme attention by researchers and professionals to minimize e-waste and use of non-toxic materials in preparation of e-equipment. Energy efficiency is increasingly important for future as the cost, and availability of energy is rising day by day.

Global warming and climate change are causing the increase of global temperature and the rise of sea levels. The main cause of environmental impact is people and their irresponsible and harmful behaviour. As an example, for such behaviours are the huge amounts of CO2 emissions from the industries and vehicles, cutting trees, and the exhaustive use of resources by technology. Studies have shown that the amount of CO2 emissions has been increasing in the past few years. Over the past few decades there has been lot of change across in temperature and weather patterns due to increase in greenhouse gases on account of massive deforestation, burning of fossil fuel and rapid industrialisation. Consequently, the average temperature of the air and the ocean has increased. Due to increase in the air temperature, there has been more melting of snow resulting in an increase in sea levels.

II. History and Evolution of Green Computing

The first appearance of green computing was the launch of energy star program in 1992 by U.S Environment Protection Agency under the authority of the Clean Air Act Section 103. Energy star is an international standard label which is designed to encourage the development of computing products with maximum efficiency and low power usage. Energy star label is provided on computer monitors, refrigerators, television sets, air conditioners and other technologies.

After this other popular green technology group came into existence that is tactical instrumentalists whose main objective is to reduce the costs rather than to protect the atmosphere.

Another major event in the history of green computing is Kyoto protocol introduced in 1997. It is an international agreement for united nation framework convention on climate change. Kyoto protocol enables the computer developers to determine the electricity used by the system over its lifetime to estimate the carbon dioxide release in order to reduce it by various preventive actions.

Another landmark in green computing history is the adoption of RoHS that is Restriction of hazardous substance and electrical and electronic equipment directive in February 2003 which prohibits the use of hazardous substances like lead, mercury, hexavalent chromium, cadmium, and polybrominated diphenyl in the development of electrical and electronic devices.

EPEAT (Electronic Product Environmental Assessment Tool) is introduced in 2005 whose main objective is to focus on various issues related to the effect of a product on environment.

A. Energy Star

Energy Star is an international standard for energy efficient consumer products originated in the United States of America.It was first created as a United States government program during the early 1990s. Devices carrying the Energy Star logo, such as computer products and peripherals, kitchen appliances, buildings and other products, generally use 20%–30% less energy than required by federal standards.Energy Star specifications differ with each item, and are set by either the Environmental Protection Agency or the Department of Energy.The following highlights product and specification information available on the Energy Star website.Energy Star items: Computers, Servers , Appliances, Heating and cooling systems , Home electronics, Imaging equipment.

III. Efforts To Implement Green Computing

We do not need to stop using computer system and power to save our environment but we have to make some effective efforts by adapting green technology to promote an eco-friendly computing environment at low cost by reducing power consumption.

- Purchase energy star labelled products
- Unplug the electronics appliances when not in use
- Use flat screen monitors rather than CRT monitors
- Use soy ink or non-petroleum-based inks for printing
- Purchase eco-friendly printing papers
- Avoid using screen savers
- E-waste management
- Recycling

IV. Benefits of Implementing Green Computing

- Green environment
- Cost saving
- Energy saving
- Secure future
- Social responsibility

V. To promote green computing the following four complementary approaches are deployed

A. Green Use

Reducing the energy consumption of computers and other information systems as well as using them in an environmentally sound manner.

B. Green Disposal

Refurbishing and reusing old computers and recycling unwanted computers and other electronic equipment.

C. Green Design

Designing energy efficient and environmentally sound components, computers, servers and cooling equipments.

D. Green Manufacturing

Manufacturing electronic components, computers and other associated sub systems with minimal impact or no impact on the environment.

VI. Overview of Existing Works on Intelligent Systems Applied to Green Computing

A hybrid combination of genetic algorithm and artificial neural network (GANN) is applied to the optimization of the energy in a wireless network toward achieving green computing. They focused mainly on the minimization of energy consumption, which was considered the objective function of a genetic algorithm. An energy savings of 51% was reported, which demonstrated the advantages of the proposed hybrid G-ANN approach.

A good method for green cloud computing based on a genetic algorithm approach is proposed. In this research work, a hybrid efficient energy algorithm based on genetic algorithms is presented. Simulation results clearly show that the proposed algorithm obtains lower energy consumption and reduces violations with high throughput compared to the best existing known algorithm.

An approach for green computing process optimization using a machine learning algorithm is presented. Handling information is an important task in the healthcare sector; for this reason, data mining techniques are the correct alternative to solve the complex problems in this sector.

In this research work, the approach of an effective recognition method is proposed for analysing heart diseases using a hybrid combination of K-nearest neighbours clustering and spiral optimization for risk classification of cardiovascular diseases. The experimental results are extremely promising, and the method can become a helpful tool for medical doctors specializing in heart diseases.

VII. Scientometric Analysis of Green Computing and Computational Intelligence Techniques

In this section, the scientometric analysis for the existing published papers in the green computing and computational intelligence areas performed for this work is presented. We present a sequence of plots and tables that show the analysis of the data (published papers) from different points of view or facets used in the analysis. We can say that this analysis helps us understand the nature of the research that has been done in the areas of green computing and computational intelligence.

Various scientometric tools were employed in this research study to understand the patterns of growth in this field.

First, we start with a co-authorship analysis of the data. Co-authorship is essential for any scientometric study, and, in the case of the interaction of green computing and computational intelligence techniques, the co-authorship can be visualized. It provides a view of the collaborative work that has been done in green computing and computational intelligence.

We illustrate the general trend on the number of publications per year since 2001, which shows an increasing trend that we believe will continue in the next years. The recent importance of environmental issues and computational intelligence applications for daily activities and industry clearly justifies the increasing trend that is shown in Figure and that we believe will continue in future years.

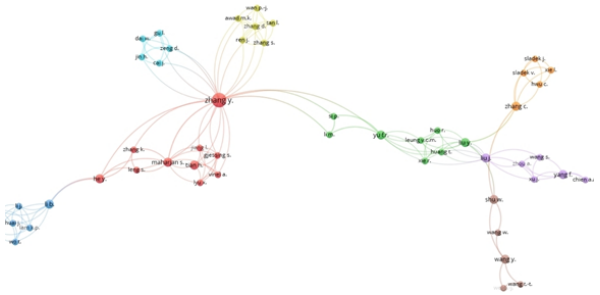
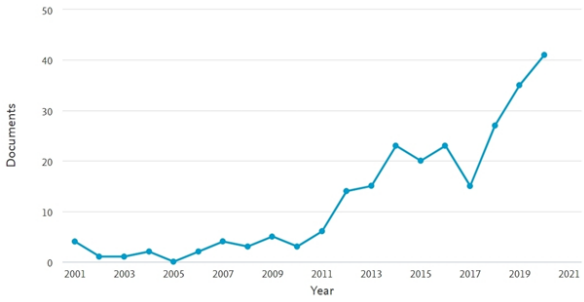


Fig. 1. Co-authorship Pattern for Green Computing and Computational Intelligence



VIII. Green Computing: From Current to Future Trends

A. Current trends of Green Computing

Current trends of Green Computing are towards efficient utilization of resources. Energy is considered as the main resource and the carbon footprints are considered the major threads to environment. Therefore, the emphasis is to reduce the energy utilization & carbon footprints and increase the performance of Computing. There are several areas:

1) Energy Consumption: Organizations are realizing that the source and amount of their energy consumption significantly contributes to Greenhouse Gas emissions. In response to this finding, organizations are currently using the following equation: Reduced energy consumption, Reduced greenhouse gas emissions, Reduced operational costs for the data center.

2) Data Center Consolidation & Optimization: Currently much of the emphasis of Green Computing area is on Data Centers, as the Data Centers are known for their energy hunger and wasteful energy consumptions.

3) Virtualization: One of the main trends of Green Computing is virtualization of computer resources. Abstraction of computer resources, such as the running two or more logical computer systems on one set of physical hardware is called virtualization.

4) IT Products and Eco-labeling: Another approach to promote Green Computing and save environment is to introduce policies all around the World, so that, companies design products to receive the eco-label

B. Future Trends of Green Computing

Future Trends of Green Computing ,the future of Green Computing is going to be based on efficiency, rather than reduction in consumption . The primarily focus of Green IT is in the organization's self interest in energy cost reduction, at Data Centers and at desktops, and the result of which is the corresponding reduction in carbon generation. The secondary focus of Green IT needs to focus beyond energy use in the Data Center and the focus should be on innovation and improving alignment with overall corporate social responsibility efforts. Some trends are:

1) Certifications: There are several organizations providing certificates to green technology. Vendors are based on their product quality, material, life of the product and recycling capabilities.

2) Cloud Computing: Cloud Computing has recently received significant attention, as a promising approach for delivering Information and Communication Technology services by improving the utilization of Data Center resources. Cloud Computing results in better resource utilization, which is good for the sustainability movement for green technology.

3) Product Longevity: the product durability and/or longevity are one of the best approaches towards achieving Green Computing objectives. Long life of product will allow more utilization of products and it will put a control on unnecessary manufacturing of products.

4) Power Management Tool: Power management is proving to be one of the most valuable and clear-cut techniques in near future to decrease energy consumption.

5) Leveraging Unused Computer Resource: One of the exiting areas where Green Computing can grow is the share and use efficiently the unused resources on idle computers.

6) Data Compression: Intelligent compression techniques can be used to compress the data and eliminate duplicates help in cutting the data storage requirements.

7) Applications : The current main applications of Green Computing are covering following computing sectors . Equipment design, Equipment recycling, Data Center optimization and consolidation, Virtualization, Paper free environment, Application Architecture and Power Management.

To be conclude Technology is not a passive observer, but it is an active contributor in achieving the goals of Green Computing. IT industry is putting efforts in all its sectors to achieve Green computing. Equipment recycling, reduction of paper usage, virtualization, cloud computing, power management, Green manufacturing are the key initiatives towards Green computing.The future of Green Computing will be depending on efficiency and Green products.

IX. Issues Related to Green Computing

In this section a brief discussion is made on various issues related to green computing.

A. Earth and the Environment

Over the past few decades there has been lot of change across in temperature and weather patterns due to increase in greenhouse gases on account of massive deforestation, burning of fossil fuel and rapid industrialisation.

B. Impact of Information Technology upon Environment

The combined effect of the energy needed to run these devices and the electricity required to maintain the cooling infrastructure for these devices have an impact on the environment.

C. Green Information Technology

Green computing should address environmental sustainability primarily by focussing on design, manufacture, use and disposal of computer and other related devices in an eco friendly way.

Green I.T 1.0 is the greening of I.T by laying emphasis on reengineering I.T products and processes to improve the energy efficiency of I.T and minimization of their carbon footprint and their environmental impact. Green I.T 2.0 is the greening by I.T and includes coordinating, reengineering and optimization of the supply chain, manufacturing process and workflow of the organization to reduce the impact on environment.

X. Green Computing on Environmental Impact

Global warming and climate change are causing the increase of global temperature and the rise of sea levels. The main cause of environmental impact is people and their irresponsible and harmful behaviour. Citizens of planet Earth must start thinking "Green" in all aspects of their lives in order to save and protect their future and the future of their children on this planet. Nowadays, technology has entered people's lives intensively - it reached their jobs, homes, and education. Efficient software reduces the use of hardware resources; therefore, reducing energy consumption. we study green approaches in software engineering models, cloud computing, mobile development, data centers.

Many efforts were done in the literature with the aim of achieving green computing in different domains and reducing the negative impacts of ICT on environmental sustainability. Berkhout and Hertin studied the impacts caused by Information and Communication Technologies (ICTs) on the environment. They presented a summary of the literature on the topic and classified the environmental impacts of ICTs into three categories as follows:

A. First-Order Impacts

the environmental impacts that result directly from the production and use of ICT infrastructure and devices such as resource use and pollution, electricity consumption of ICT hardware, and disposal of electronic waste.

B. Second-Order Impacts

indirect environmental impacts of using ICT such as resource and energy conservation caused by dematerialization, demobilization and substitution of information goods for tangible goods.

C. Third-Order Impacts

indirect environmental impacts of using ICT that appears in the long term such as changing lifestyles and values systems. These impacts may overcompensate for the energy savings by ICT.

The GREENSOFT model also supports different stakeholders of a software product in developing, maintaining, and using it in a sustainable manner. The model, four parts are Life Cycle of Software Products , Sustainability Criteria and Metrics , Procedure Models and Recommendations and Tools.

The first threat consists of the inability to manage the cost crises that are born due to the divergence in determining the efficiency performance technique and calculating the performance of the server. Another way to produce efficient energy that was provided in this study is either the use of virtualization software through a virtual machine or avoids the rapid change between alternating and direct current. They managed to divide their area of work into four fields by presenting energy-saving techniques for servers, energy-saving techniques for networks, energy saving techniques for a combined environment of servers and networks, and finally by energy-saving techniques by using renewable energy.

XI. Challenges of Green Computing

In the past the focus was on computing efficiency and cost associated to IT equipment and infrastructure services were considered low cost and available. Now infrastructure is becoming the bottleneck in IT environments and the reason for this shift is due to growing computing needs, energy cost and global warming. This shift is a great challenge for IT industry. Green Computing challenges are not only for IT equipments users but also for the IT equipments Vendors. Several major vendors have made considerable progress in this area. Major companies like HP, DELL, IBM are working on technologies to improve the more energy and cost efficient systems.

XII. Conclusion

Technology is not a passive observer, but it is an active contributor in achieving the goals of Green Computing. IT industry is putting efforts in all its sectors to achieve Green computing. Equipment recycling, reduction of paper usage, virtualization, cloud computing, power management, Green manufacturing are the key initiatives towards Green computing. Green computing has gain lot of importance due to rise in power consumption and its impact on environment.

Green Computing following are few prominent challenges that Green computing is facing today:

- Equipment power density / Power and cooling capacities
- Increase in energy requirements for Data Centers and growing energy cost
- Control on increasing requirements of heat removing equipment, which increases because of increase in total power consumption by IT equipments
- Equipment Life cycle management – Cradle to Grave and Disposal of Electronic Wastes.

Going green is very beneficial as it helps various industries to manage their E-waste in an efficient way so that the surrounding may not be effected. There are still other kinds of models coming from the computational intelligence area that can be applied to green computing problems. The computational intelligence area is also evolving in new directions, such as the use of quantum computing, nanotechnology, and other newer technologies that may come to be developed in conjunction with intelligent techniques and models.

References

[1] Robert R. Harmon, Haluk Demirkan, The Corporate Sustainability Dimensions of Service-Oriented Information Technology, Annual SRII Global Conference, March 29 2011 – April 2 2011, DOI: 10.1109/SRII.2011.116

[2] PatrikKurp, Green Computing – Are you ready for a personal energy meter?, Communication of the ACM, 2008, Vol 51, No. 10, DOI: 10.1145/1400181.1400186

[3] Climate Savers Computing Initiative (2010) Retrieved

[4] Ryan, John C. & Durning, Alan T. Stuff: The SecretLives of Everyday Things. 1997

[5] Benotmane Z, Belalem G, Neki A. Towards a cloud computing in the service of green logistics. IJLSM 2018;29:37. DOI

[6] Baliga J, Ayre RWA, Hinton K, Tucker RS. Green cloud computing: balancing energy in processing, storage, and transport. Proc IEEE 2011;99:149-67. DOI

[7] Naumann, S., Dick, M., Kern, E., Johann, T. (2011). The GREENSOFT model: a reference model for green and sustainable software and its engineering, Sustain.

[8] Berkhout, F. and Hertin, J. (2001). Impacts of Information and Communication Technologies on Environmental Sustainability: Speculations and Evidence, Report to the OECD

[9] Mahmoud, S.S., Ahmad, I. (2013). A green model for sustainable software engineering.

[10] Userful, Userful is the Green Solution: reduce CO@ emissions and electronic waste, 2011

[11] Andreas Berl, Erol Gelenbe, Marco Di Girolamo, Giovanni Giuliani, The Computer Journal, 2009, Volume 53, Issue 7, pp. 1045-1051, DOI: 10.1093/comjnl/bxp080

Cloud Computing Architecture

Amritha Suresh, Diyana Ummer and Josiya Mary K S
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Aparna S Balan
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: aparna@vidyaacademy.ac.in)

***Abstract*—The Cloud has gained large popularity as means for saving cost of IT ownership , dynamically scalable computing infrastructure and software services offered on Cloud on pay-per-use basis. Cloud Computing is evolving and has always been an a key computing platform for sharing resources that include infrastructures, software, applications, operation procedures and business process. As cloud computing technology is being used to minimize the usage cost of computing resources, many enterprises gained interest of migrating their old system to the cloud computing system. Currently, there is only a few cloud computing architecture that can be used as a reference for building a cloud computing infrastructure and this paper is an attempt to discuss few among them. Cloud computing architecture is used as a guideline to understand the whole process including actor roles inside a cloud computing environment. This paper present an overview of nine cloud computing architecture and its components focusing on various computing aspects such cloud service management, travel reservations solution, Platform as a service (PaaS) infrastructure , by bridging the power of SOA and virtualization, Quality of Experience(QoE), etc.**

***Index Terms*—Cloud computing architecture,SOA(Service oriented architecture).**

I. Introduction

As the information age goes faster, the users need for various resources and needed efficiency are increasing day by day and in the reference cloud computing has gained large popularity. Cloud computing is the on demand delivery of computing services such as storage, database, computing power etc over the internet i.e, the users can access technology services with pay-as-you-go pricing without owning, buying and maintaining physical data centers and servers. By past years cloud computing has become trending technology and has shown rapid adoption as corporates/enterprises switched their services on the cloud because of its wide range of benefits such as improved collaboration, excellent accessibility, low maintenance cost, mobility, unlimited storage capacity etc. The cloud computing has always been an important computing platform for sharing various resources including infrastructures, software, operation procedures and application programs. Cloud offers an almost endless list of services that are constantly being expanded. Some popular cloud services are IaaS [Infrastructure-as-a-Service], PaaS [Platform-as-a-Service] and SaaS [Infrastructure-as-a-Service]. In the whole cloud computing, the Platform as-a-Service, also called sharing middleware takes couple steps further taking care of physical hardware, operating system and development kit.

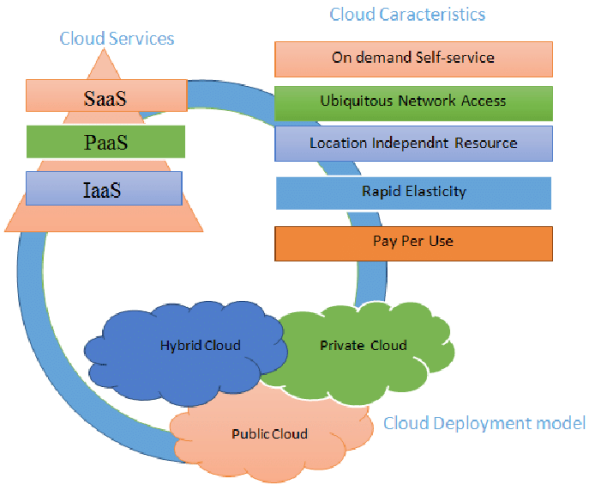


Fig. 1. Types of Cloud

Cloud computing is a kind of service model, which provide services of different types and levels adapt different methods according to different services. Cloud computing also deals with Business Information Systems BIS) scenario requirements, such as Customer Relationship Management (CRM) or Enterprise Resource Planning (ERP) systems, which require the capability for extending Data Models inside the MTA and also provide more knowledge oriented metadata-driven approaches such as inference, where data relationships can be entailed. Cloud computing offers the potential to dramatically reduce the cost of software services through the commoditization of IT assets and on demand usage patterns. However, the complexity of determining resource provision policies for applications in such complex environments introduces significant inefficiencies and has driven the emergence of a new class of infrastructure.

On the other hand the video streaming services are responsible for the majority of the Internet traffic. A good cloud-level architecture partially solves some issues related to the video streaming services. However, it introduces other issues such as

higher latency and core network congestion, addressed in this work. Video streaming services have stringent requirements such as a good-quality communication channel as well as a steady and uninterrupted flow of information. Because of that, the employment of such services in a fog/cloud environments have attractive advantages to improve the end users Quality of Experience (QoE). Examples of services include video transcoding, multiple route video transmission, and cache schemes. A transcoder service can be used to transcode a video with a bit rate of 8 Mbps (1080p) to 5 Mbps (720p), with no visible loss in quality, if the end-user device is not ready to display videos in 1080p. This allows a larger number of users to be served while maintaining a satisfactory QoE. Moreover, smartly caching videos in edge nodes can reduce traffic load and delay, since multimedia content may be readily available closer to the users. In hybrid cloud approach, raw sensor data is stored on user's personal cloud so that the users will have a feeling that the data is under control with full knowledge of where the data is stored and managed and will thus have less privacy-anxiety.

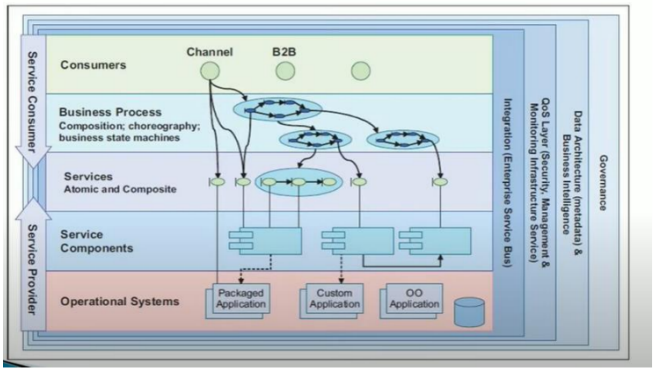
Several organizations (eg: NIST and IBM) have developed various cloud computing architecture based on above mentioned aspects and have contributed for further research or development in cloud computing. Cloud computing architecture states components and subcomponents embedded in it. Understanding or just knowing cloud computing architecture is not good enough for building a cloud computing environment or even developing cloud services. Companies should consider having several businesses and IT expertise that are expert and professional on their field to help them understand what requirements are required to become a successful service provider.

II. Cloud Computing Middleware Architecture

In whole cloud computing ,among all cloud computing services the Platform as-a-Service also called sharing middleware takes couple steps further taking care of physical hardware, operating system and development kit. If PaaS is the core of the cloud computing system, then the middleware [intermediate zone of the system] is the core of PaaS. The key point of middleware is to form middleware resource pool and the application container, it manages the computing resources ,the network communication and it offers exploitation and operation environment for internet applications. In addition it provides communication and exchange services for various businesses to solve the interconnection problem between systems. The middleware includes the application middleware and the integration middleware[data integration and application integration].Cloud computing is a kind of service model. Service Oriented Architecture[1] can be regarded as a method to provide the service integration for cloud computing. SOA is a component module, which sends the application program as service to the users or other services, and connect them through good interfaces.The core factor of SOA include:

A) Standard packaging(interoperability):
   SOA defines a way to make software components reusable and interoperable via service interfaces. SOA achieve interoperability through series of standards to achieve access, link and semantic interoperability. SOA achieves the link interoperability through standard SOAP protocol and depends on special access protocol such as JAVA use RMI . In addition, the service packaging also adopts XML protocol processing self analysing and self definition.
B) Software reuse:
   Reusability of software component is a key determinant of application development and maintainability of the application. The reuse of SOA solve the problem of faced during earlier reuse technology by subprogram. The components such as the DLL components invented under MS operation system can be reused within one system /same OS, and is dynamic during operation. SOA takes the services as the core services such as WebService, SCA/SDO and through these SOA achieve higher level of reuse, decoupling, and inter-operation which is also called the SOA structure middleware.
C) The Coupling Relationship:
   Loose coupling is a fundamental concept of SOA(applies to large distributed system) employed to deal with the requirements of scalability, flexibility and reduce dependencies. The traditional software couples sections such as operation logic, internet link and data transmission into integration and SOA structure achieves the complete decoupling of these section through service packaging.

Service oriented architecture is a design pattern that is designed to build distributed systems that deliver services to other application through the protocol. It defines a way to make software components reusable and interoperable through service interfaces. One of the key feature of SOA is the use of loosely coupled connection among services or applications. This loose coupling aspect of SOA benefits the business environment as it allows to adapt to changing business needs on demand, i.e, loose coupling in SOA offers flexibility.Service modelling is the first step, which is fundamentally a process of service identification and granularity confirmation. Service identification is one of the main key activity in modelling of a service oriented solution and various strategies exist to identify the list of the candidates which can turn to be service within certain range, the connector of the service. Service granularity is an important key feature of SOA that impacts other internal SOA quality attributes. Service encapsulating period decide in what form the service is provided for outsiders. Service governance in SOA is set of processes used to exercise control over services in SOA. It effectively manage the already encapsulated services. It provides service registration, safety controls, SOA testing, version management etc.. Service arrangement and Service assembly of SOA realize the business process and arrange and assemble the services accordingly. Application delivery period is the final phase to realize the business on demand delivery and service is ready to deploy.



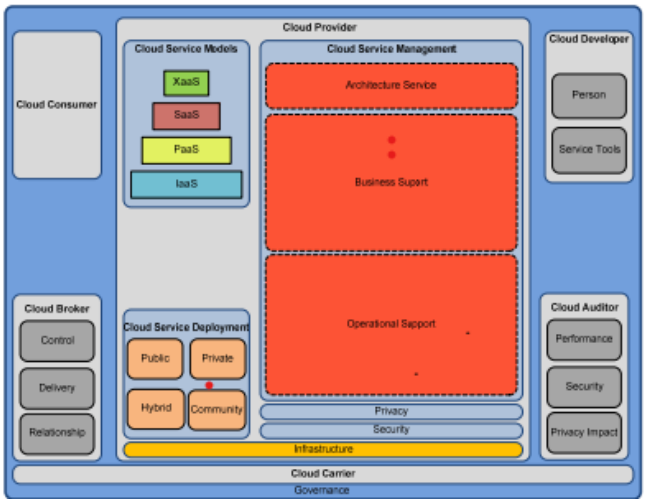
III. Cloud Computing Reference Architecture: Cloud Service Management Perspective

The cloud reference architecture[2] is an actor based model which allows people to understand more deeply about components and several activities and focus on cloud service management(cloud provider component). This CCRA has six major actors and governance :

A) Actors: Cloud actors are basically cloud players. Each cloud actors has its own Activities, requirements, and responsibilities. There are six main actors in Cloud computing.
   a. Cloud consumers: Cloud consumers are the end users. According to the cloud consumers Expectation, they browses or utilize services provided by cloud service Providers. Based on cloud service model, cloud consumers can be Classified to four different groups–IaaS (Infrastructure as a Service) Consumer, PaaS (Platform as a Service) consumer, SaaS (Software as a Service) consumer, XaaS (Anything as a Service) consumer.
   b. Cloud Providers: Cloud providers provide services to cloud consumers and manage cloud Computing environment. Once the services are delivered and is used by Cloud consumers, cloud providers manage, maintain, monitor and operate The process. In addition, cloud providers must understand the privacy, Security requirements, cloud service models, cloud service deployement Etc..
   c. Cloud developer: Cloud developers are person or application who create applications that are Served on the cloud. They are responsible for developing and implementing Cloud services. Basically, they offers cloud services to cloud providers or Deploy it in the cloud providers system.
   d. Cloud broker: Cloud broker is an entity that negotiates relationship between cloud Consumers and providers. It manages the performance, use and delivery of Services.
   e. Cloud Auditor: Cloud auditor is a third party who perform an audit to improve cloud Service quality by performing evaluation, assessment, and examination in Terms of performance, security, and privacy of cloud services.
   f. Cloud Carrier: Cloud carrier is an intermediary that provide connectivity and Transportation of cloud services from cloud providers to the consumers.

B) Governance: Governance in cloud computing comprises of set of rules, policies and Regulations to manage, control and maintain cloud services to ensure its better Quality and performance. It defines how a service should be done and by Whom the activities and processes should be performed.

Cloud service management is an important component of cloud providers in the Above depicted architecture. Cloud service management provides necessary Functions, methods, activities for managing and operating cloud service for the Customers. Cloud service management[3] is supported by three major support Services -architecture service, business support, and operational support. Architecture service includes architecture design and architecture deployment That supports the consulting process of the cloud architecture integration and Alignment to cloud customer business architecture and strategy. Business support Interacts with business-related services that directly deal with the cloud Consumers. Operational support deals directly with the service level agreement Management and the monitoring of the cloud service operations to ensure that the Service level agreed can be delivered to the consumer



IV. Cloud Computing Open Architecture

This Cloud Computing Open Architecture is based on seven architectural principles and ten architectural modules, by integrating the power of serviceoriented architecture (SOA) and virtualization technology of hardware and software. Virtualisation allows to run multiple operating system and application to run on same server by creating a virtual environment. In addition, virtualisation allows subscribers to reuse licences of

operating system, middleware or software application once he/she releases services from cloud computing platform. SOA is a design pattern that provides a way to address reusability, extensibility, flexibility and interoperability of software components through software interface. This cloud computing open architecture (CCOA) is an integrated co-production framework where cloud actors work together based on seven principles which is use to define ten main architectural modules as depicted in the figure. The Cloud CCOA is an extensible and configurable architecture that covers cloud ecosystem enablement, cloud infrastructure and its management, service-orientation, cloud core on provisioning and subscription, compostable cloud offerings, cloud information architecture and management, and cloud quality analytics.The seven principles of CCOA are as follows:

A. Integrated Ecosystem Management for Cloud:
The cloud computing ecosystem includes all involved services and solutions vendors, partners, and end users to provide or consumer shared resources in the Cloud Computing environment. Architecture must manage this cloud computing ecosystem. From the perspective of this architecture, the vendor dashboard provides a integrated view of interaction of vendors frond end and back end. In most cases cloud vendors collaborate with their partners in cloud computing environment and in this regard a partner dashboard is needed for the collaboration. The cloud client dashboard provides a framework for clients/end users to interact with or consume cloud services via multiple channels such as Web portal, program-based business to business collaboration channel, or phone based customer representative channel. Putting all these dashboards collectively, the cloud computing ecosystem management layer (1A) provides an integrated on-boarding process and common utilities to support the seamless collaboration and message exchanges among cloud vendors, partners, and clients.

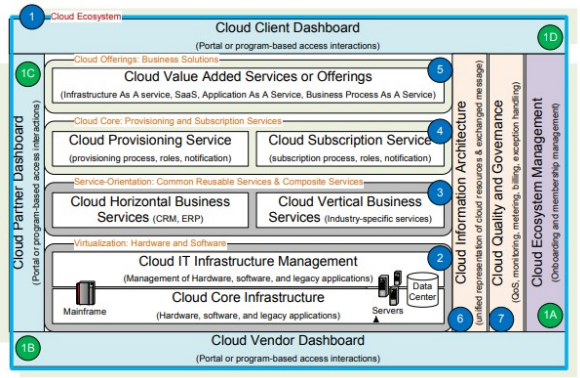
B. Virtualization for Cloud Infrastructure:
Virtualisation is a fundamental technology of cloud computing. The two basic approaches to enable virtualization in cloud computing are hardware virtualization and software virtualization. Hardware virtualisation allows virtual machine software or VMM to directly install on the hardware system. Also, hardware equipments can be added or removed without affecting the normal operations of other equipments in the system. The software virtualization use software image management or software code virtualization (dynamic code assembly or execution) technology to enable software sharing. The Cloud IT Infrastructure Management module covers software image management, hardware virtualization, and legacy application packaging(home-grown applications or ISV applications).

C. Service-Orientation for Common Reusable Services:
Service orientation is another core driving force to enable cloud computing which aims to enable reusability, interoperability, composition to realize the business value. Cloud Horizontal and Vertical Business Services are the two major reusable services. The Cloud Horizontal Business Services offers platform services , middleware, some common utilities such as onboarding, provisioning, monitoring etc.. The Cloud Vertical Business Services comprises of all domain specific or industry-specific utility services such as shipping and payment services.

D. Extensible Provisioning and Subscription for Cloud:
Cloud provisioning is a key feature of cloud computing which allows allocation of services and resources to the customers by the providers. Without extensibility, the provisioning of cloud computing architecture only supports sharing of certain resources. As depicted in the figure, the cloud core is divided in to Cloud Provisioning and Subscription Service. Cloud Provisioning Service include provisioning process, role definitions, and notification framework whereas Cloud Subscription Service involves subscription process, role definitions, and notification framework.

E. Configurable Enablement for Cloud Offerings:
Cloud offerings are the final products or computing services such as storage ,servers, software, databases etc.. that are provisioned by the Cloud Computing platform. CCOA defines its offering aspects at the following four levels- infrastructure as a service, software as a service (SaaS), application as a service, and business process as a service. In the category of infrastructure as a service, CCOA uses the cloud core to facilitate provisioning and subscription of virtualized IT infrastructure resources, taking advantage of some common reusable services and cloud ecosystem management features. In the area of software as a services, many success stories demonstrate common reusable services like CRM as a service (e.g., Salesforce.com), payment as a service (e.g., eBay's PayPal), and shipping as a service by exploiting the cloud core to manage provisioning and subscription. In the application as a service area, lots of standalone applications or network-based applications can be provisioned as cloud services for sharing in the Cloud Computing value chain. Business process as a service is a new concept for sharing best practises and business processes among cloud clients and value chain partners in the Cloud Computing environment.

F. Unified Information Representation and Exchange Framework:
To enable the collaborative and successful aspects of Cloud Computing, information representation and message exchange of Cloud Computing resources are critical. The cloud information architecture module in CCOA represent cloud entities, both business entities (cloud clients, partners, and vendors) and supporting resources (virtualization related modules, serviceorientation related modules, cloud core, and cloud offerings.) in a unified Cloud Computing entity description framework.


Cloud Ecosystem
Cloud Client Dashboard
(Portal or program-based access interactions)
Cloud Offerings: Business Solutions
Cloud Value Added Services or Offerings
(Infrastructure As A service, SaaS, Application As A Service, Business Process As A Service)
Cloud Core: Provisioning and Subscription Services
Cloud Provisioning Service
(provisioning process, roles, notification)
Cloud Subscription Service
(subscription process, roles, notification)
Service-Orientation: Common Reusable Services & Composite Services
Cloud Horizontal Business Services (CRM, ERP)
Cloud Vertical Business Services (Industry-specific services)
Virtualization: Hardware and Software
Cloud IT Infrastructure Management
(Management of Hardware, software, and legacy applications)
Cloud Core Infrastructure
(Hardware, software, and legacy applications)
Mainframe
Servers
Data Center
Cloud Partner Dashboard
(Portal or program-based access interactions)
Cloud Information Architecture
(unified representation of cloud resources & exchanged message)
Cloud Quality and Governance
(QoS, monitoring, metering, billing, exception handling)
Cloud Ecosystem Management
(Onboarding and membership management)
Cloud Vendor Dashboard
(Portal or program-based access interactions)

G. Cloud Quality and Governance:
From quality indicators' perspective, QoS parameters can be directly used to define reliability, response time, security, and integrity of cloud entities. Security is of prime importance and Access control, privacy, and protection of the cloud entities form the aspects of the trust and security foundation in CCOA.SOA governance consists of many best practices that can be adapted to enable Cloud Computing environments and to govern the design, deployment, operation, and management of the cloud services offerings.

V. CLOUDIO: A CLOUD COMPUTING-ORIENTED MULTI-TENANT ARCHITECTURE FOR BUSINESS INFORMATION SYSTEMS

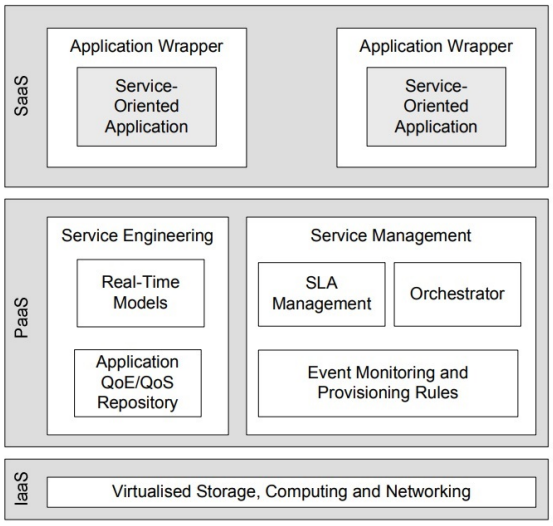
When dealing with Business Information Systems BIS) scenario requirements, such as Customer Relationship Management (CRM) or Enterprise Resource Planning (ERP) systems, which require the capability for extending Data Models inside the MTA and also provide more knowledge oriented metadata-driven approaches such as inference, where data relationships can be entailed. Cloud technologies are seen more as a new business model, rather than a new technology. Multi-tenant architecture is a core technology that enables cloud computing. Multi-tenant is used by all public cloud providers, and is universally deployed in the cloud landscape. Indeed, multi-tenant architecture literally makes cloud computing economically and technologically feasible by allowing a mixed number of customers to leverage one platform.The architecture of this system has been designed to offer the best flexibility and usability for the multi-tenancy database. The major contribution of this paper is the two levels of multi-tenancy that this architecture provides. First level of multi tenancy is due to different tenants are assigned to different databases (Different Instances), and the second one is provoked for the ability of the tenants to share the tables of the database (Shared Schema). It has been possible adding two innovative elements to the typical architecture: the "JDBC Driver Manager" and the "Multi Tenant DB Index", which together make the transition from single-tenant to multi-tenant almost automatic in a very flexible way. These two artifacts make this architecture different from a regular multi-tenant application working in the Amazon or another cloud. The SaaS application is in the top of the architecture and it is a J2EE application that needs to store information in a database. Actually, it does not need to be a "SaaS" application but this paper is focused on SaaS applications because multi-tenancy is a key feature in the Cloud Computing world. The JDBC Driver Manager is just above the pre-defined JDBC driver and allows, as it was explained, a multi-tenant database to be created automatically with no effort in the tenant side. In the same layer there is the Multi-Tenant DB Index that contains the data of which database is the correspondent for each tenant, therefore it is implemented as a hash map of key value pairs, being the key the tenant id, and the value the connection parameters and the subjacent JDBC driver to use. Under the JDBC Driver Manager it is located the JDBC works.The implementation consists of two parts: the necessary modifications to the database schema, and the driver implementation.

A. Database modification:
First, a new table is created to store tenant data, with a single field, the tenant id. This table is not strictly needed, but it allows the user to associate data with tenants (e.g: tenant name) and to maintain data consistency in the database, forcing every data to be owned by one of the tenants on this table. The other bit of data that is added to the database is a new column, storing the identifier of the data owner, in every existing table.

B. Driver implementation:
The actual driver is implemented as a JDBC driver, so any Java application can use it without significant modifications. When a Java program requests a new connection, the driver starts a real connection to the database using a second database driver. The back-end driver to use, as well as its connection parameters, depends on the tenant that is requesting the connection to the database, which is passed from the application to the front-end JDBC driver.

VI. PLATFORM-AS-A-SERVICE ARCHITECTURE FOR REAL-TIME QUALITY OF SERVICE MANAGEMENT IN CLOUD

Cloud computing offers the potential to dramatically reduce the cost of software services through the commoditization of IT assets and on demand usage patterns. [4] Investigate the problem of shared resource contention in virtual machines and propose a model to estimate the shared resources required. One important characteristic of the virtual machines proposed by the authors is their capability to enforce policies to guarantee QoS parameters in SLA. However, the complexity of determining resource provision policies for applications in such complex environments introduces significant inefficiencies and has driven the emergence of a new class of infrastructure called Platform-as-a-Service A PaaS is built on IaaS and uses both a multi-tenanted development tools and deployment architecture.



A good example of PaaS is Facebook – a venue where multiple applications share resources and user information, subject to tight controls. (PaaS).

a. The PaaS hoster:
whether external or internal, mustprovide adequate resources (typically via an IaaS model) in order to meet demands of its customers needs, together with appropriate availability contingencies.
b. The PaaS provider:
: whether an external or internal provider will provide an environment suitable for general developers to build web applications without deep domain expertise of back-end server and front end client development or website administraion.
c. The PaaS user(developer):
Must have a browser based development environment; the ability to deployment seamlessly to a hosted runtime environment; management and monitoring tools; and pay as you go billing.

A PaaS architecture supporting real time interaction between distributed set of people and resources requires the following key features:

a. Real-Time QoS Specification :
specification language and associated toolkit for the specification of IRMOS applications and application service components considering both structure and real-timeQoS
b. Event Prediction:
QoS oriented service engineering methodology and models for predicting QoS requirements contingent on application and resourcing events considering temporal profiles of application service components deployed on virtualised infrastructures.
c. Dynamic SLA Negotiation:
SLA negotiation and management services supporting the dynamic negotiation (self-service) of Application-SLAs considering customer requirements and dynamic discovery of resource providers (Technical-SLAs).
d. On-Demand Resource Provisioning:
Provisioning services for application service components on virtualised infrastructures through combination of workflow and service-based management wrappers enhanced to support temporal profiles
e. QoS Event Monitoring:
Monitoring services for measuring quality of service at both application and technical levels targeting trigger events for runtime adaptability of resource provisioning estimation and decision making.

A. QoS oriented Service Engineering Quality of Service oriented Service Engineering (QoSSE) supports two important features for real systems: Real Time Specification for Applications and Event Prediction. The activities in service engineering . As previously stated, Real-Time Specification is related in how an application is represented to the PaaS architecture and Event Prediction is associated with the procedures and mechanisms needed to model an application and determine the appropriate infrastructure requirements.
B. On-Demand Service Management :
The service management components are responsible for on-demand SLA negotiation, resource reservation, service instantiation, execution and monitoring of Quality of Service for application provisioned at an IaaS provider. In IRMOS the IaaS is provided by ISONI (Intelligent Service Oriented Network Infrastructure)
C. SLA Management:
The SLA management component consists of three sub-components the SLA Negotiator, A-SLA Management and T-SLA Management. The SLA Negotiatoris responsible for SLA negotiation. It aims to provide valid SLA offer to the service customer prior to the execution of the services so that the service level requirements can be guaranteed while meeting customer's satisfaction. There are two different negotiations involved.
D. Orchestrator:
The orchestrator is a controller responsible for configuring, starting and stopping the applications. The role of the orchestrator can be two-fold: it is responsible for configuring the applications prior to execution, which is achieved by receiving a configuration command from the A-SLA manager.
E. Event Monitoring and Provisioning Rules:
The Monitoring Service is responsible for collecting the run-time information generated by applications and the IaaS respectively. Functionalities of the Monitoring Service includes aggregating and storing the acquired data to the Historical Database (for offline usage - providing input to the QoSSE), and detecting on-the-fly any A-SLA violations (online usage - providing input to the A-SLA manager).Also related to cloud computing for real-time systems is the work done by Kim, Beloglazov and Buyya

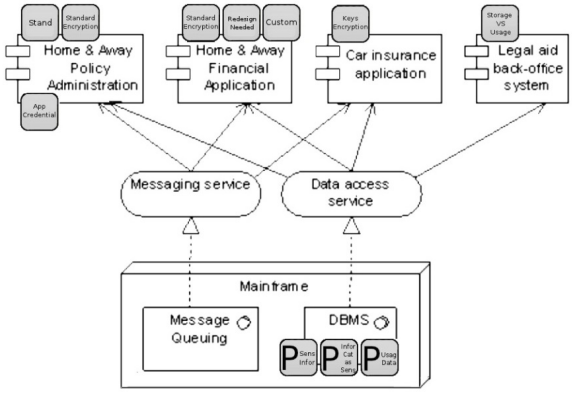
[5].

F. ASC Wrapper:
IRMOS involves different type of ASCs (e.g. transcoding, colouring), each of which has its own specific features. In order to decouple the specific ASC from the PaaS services, an ASC Wrapper is developed to 'wrap' the specific ASC binary and provide a unified management interface. The ASC wrapper itself is generic.

VII. Enterprise Architectures for Cloud Computing

Enterprise Architectures was born as an information technology discipline. Initially the purpose of Enterprise Architectures was to promote the strategic development of an organizations IT systems through the modelling of the organization,as well as the aligning of IT purchasing and development with business priorities[8]. In [6] an Enterprise Architecture is a strategic information asset base, which defines the mission,the information necessary to perform the mission and the transitional processes for implementing new technologies, in response to the changing mission needs. An enterprise architecture characterizes and models the enterprise throught a set of interrelated layers or views: strategy, business, data, applications, and technology. If Enterprise Architectures would be suitable for the evaluation of cloud readiness they would allow statements with respect to recommendations. Unfortunately Enterprise Architectures don't include any kind of information related to cloud computing issues. For example, Enterprise Architectures miss any kind of a description of service level agreement (SLA).Furthermore nearly any recommendation is hard to be determined in current Enterprise Architecture frameworks.This is the main reason why an extension to the architectural representation of the company is needed.There are many areas and related best practices which should be considered by a company that desires to move into the cloud. For example important recommendations identified deal with application security and encryption. Some recommendations concerning application security and encryption can be identified, e.g. [7] includes: 1) the management of application credentials is critical: they should be protected and secured; 2) it's important to use encryption to secure the storage of data in the cloud environment. In a similar way, data is protected by an encryption method during its usage within a cloud environment, as well as in an external system; 3) the chosen encryption system should conform to the existing industry and government standards; 4) keys used to encrypt data must themselves be protected when they are stored, in transit or backed-up; 5) check if the data are encrypted also during the transit inside the cloud provider's network. Cloud provider's network is more secure than the Internet but, anyway, it's shared between different customers; 6) it's possible that in IaaS scenario, also virtual machines files and temporary data need to be encrypted.

The graphical extension considers each level of a generic EA framework. In each layer, some icons are added each of which represents an issue related to the cloud readiness



of the company.Presence of DBMS leads to the necessity to indicate that they manage data with different levels of privacy. Concerning the applications, in the current level, they just inherit their features described in the application level. All the symbols used to mark a specific application will be reported in this level. As explained later on, applications are tagged with icons expressing their level of standardization.The main graphical extensions for application level in 2 deal with different issues. For example, considering the area of application security and encryption, each application should be marked in order to show if their level of security should be increased, in order to go into the cloud.The knowledge representation allows to close the gap between the recommendations and the graphic representation.The knowledge representation considers the best practices and describes them in a formal way, in order to automate them. The formal description used for the knowledge representation is the business rules method. The formalization of the recommendations helps to link them to the graphical representation. In more detail, it helps to link the recommendations to a specific layer (or more than one) of the general Enterprise Architecture framework used and to a specific symbol.

VIII. Service Oriented Architecture for Cloud based Travel Reservation Software as a Service

A. Service-Oriented Architecture

Service-Oriented Architecture (SOA) is a stage in the evolution of application development and/or integration. It defines a way to make software components reusable using the interfaces. Formally, SOA is an architectural approach in which applications make use of services available in the network. In this architecture, services are provided to form applications, through a network call over the internet. It uses common communication standards to speed up and streamline the service integrations in applications. Each service in SOA is a complete business function in itself. The services are published in such a way that it makes it easy for the developers

to assemble their apps using those services. Note that SOA is different from microservice architecture.

B. Architecture-Design

This section provides architecture-design details of prototype implemented by us for SaaS and Cloud based Travel Reservation Software as a Service (TRSaaS) to address the system stakeholder requirements.

- Logical View
- Deployment view

1) Logical components in the architecture

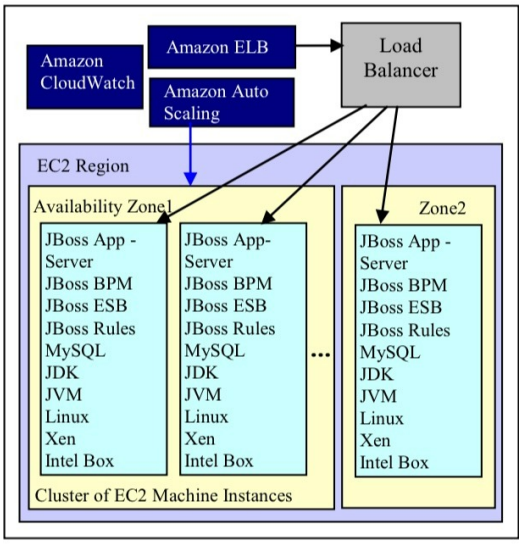


Figure 2. Deployment on Cloud

- Interfaces TRSaaS interacts with multiple tenants and multiple suppliers. It provides a Website which allows the tenant side administrators to upload their organizational master data and specify customizations needed. TRSaaS provides Metadata Services for provisioning the customizations per tenant. After the meta-data is configured, the tenant side end-users can place the reservation order by choosing airline and hotel through the Website. Airline supplier organization and hotel supplier organization expose their respective internal system functionality through Web Service interfaces. Bank application provides the funds verification Web Service interface. TRSaaS uses asynchronous email notification to inform status of the order placed to user.
- Databases TRSaaS segregates the supplier data and consumer tenant data into different databases. Additional data stores are provided for TRSaaS internal processing data, and for metadata that also serves as service registry. Data Services are used for uniform access, manipulation of data across multiple data sources.
- Scalability The software is layered, decomposed into runtime components and static modules for better scalability.
- Security Users of the system are subjected to Login-Password based Authentication and Role Based Access Control mechanisms. Communication links between TRSaaS and all on-premise applications are secured using Secured Sockets Layer (SSL).
- Service Oriented Architecture TRSaaS provides Metadata Services that for provisioning tenant specific variations, Data Access-Manipulation Services, Data Synchronization Services for synchronizing data of external parties, Monitoring and Service Level Agreement (SLA) Management Services, Charging and Reporting Services for monthly subscription plus pay-per-use basis billing, and Revenue Tracking Services. TRSaaS uses Enterprise Service Bus (ESB) middleware for hosting, integration, and lifecycle management of these services, For infrastructure level services, TRSaaS relies on Amazon Web Services, viz., Elastic Computing Cloud (EC2) service for creating Machine Images and instances of the deployment environment needed, Simple Storage Service (S3) for storing dynamic data pertaining to Amazon Machine Images, CloudWatch and Auto Scaling services for resource utilization monitoring and scaling, Elastic Load Balancing (ELB) service for workload distribution, as well as Amazon Cloud provided facilities for Security, Metering and Billing.
- Multi-tenant Architecture Single copy of the TRSaaS serves multiple tenants. Configurability required for multi-tenancy is achieved using Generalization Specialization technique used within Object Oriented modeling and through Metadata modeling for the parameters that are variable across tenants. Metadata for a tenant is used to provide customized look and feel of the Website to the tenant with corporate color scheme, fonts, and logo. The screen layouts are designed using tenantspecific activity steps are managed using the Business Process Management (BPM) solution. A default process template is provided to all tenants that can be customized template is provided to all tenants that can be customized to their specific needs. Tenants can customize the default business rules and access control through Metadata Services. TRSaaS provides multiple options for multi-tenancy at database level: Shared Database with Shared Schema, Shared Database with Separate Schemas, and Dedicated Database.

2) Deployment components in the architecture At Cloud infrastructure level, Amazon Web Services and associated functionality is used to handle technical requirements of TRSaaS.

- Security: Security requirements are met through Amazon's Cloud sites with controlled physical access. distributed storage provided by S3, default inbound firewall provided by EC2, and it add filters to restrict inbound and outbound traffic.
- Scalability: The system uses Amazon CloudWatch

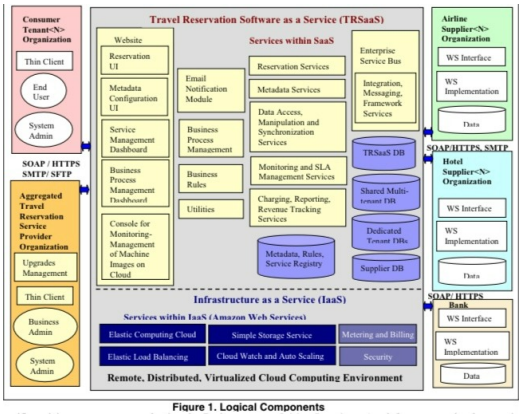


Figure 1. Logical Components

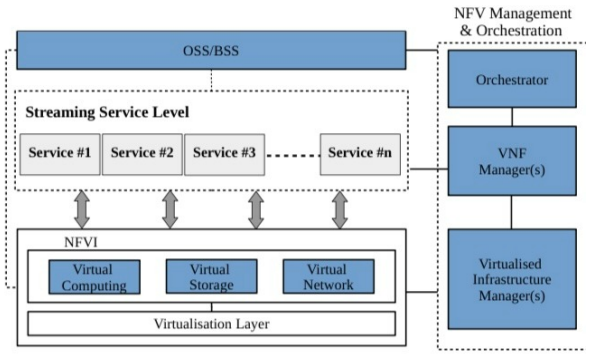
service that provides real-time visibility into resource utilization Amazon Auto Scaling service was called to address the need for scaling up or down of the EC2 capacity based on the metrics provided by CloudWatch.

- Performance: The architecture uses Elastic Load Balancing Service provided by Amazon EC2 to balance load between multiple machine instances within a zone and across zones of a geographic region defined by Amazon

IX. Towards a Multi-tier Fog/Cloud Architecture for Video Streaming

The proposed model is responsible for providing a set of multimedia services dynamically deployed in the fog/cloud network. Users can access multimedia content through a variety of heterogeneous communication technologies. In scenarios like this - common in fogs - fog nodes can provide computing and storage services[11]. The top tier is composed of cloud servers, which can be located in a public or private cloud. The other three layers represent the fog/edge network. This work takes into account edge nodes with computational and storage capacity and classifies them hierarchically according to their coverage and communication technology

A. Fog-Cloud Architecture



This work follows the framework architecture specified by the European Telecommunications Standards Institute(ETSI). This framework takes advantage of Software Defined Network (SDN) and Network Function Virtualization (NFV) principles [5]. The SDN/NFV-based approach allows virtualizing specific services on remote servers. In this way, the migration of services can be virtualized in different datacenters on the fog/cloud networks. The ETSI-NFV framework architecture includes mainly three functional modules, namely NFV Orchestrator (NFVO), VNF Manager (VNFM), and Virtual Infrastructure Manager (VIM). The ETSI-NFV framework architecture includes mainly three functional modules, namely NFV Orchestrator (NFVO), VNF Manager (VNFM), and Virtual Infrastructure Manager (VIM). The NFVO is a software component that can orchestrate the lifecycle of virtualized network functions. This includes creating, monitoring, and chaining network services. The VNFM is responsible for handling specific VNF instances, coordinating requests for infrastructure resources between the VNF instance and related network modules management systems. The VIM controls and manages the NFV infrastructure, which includes computing, storage, and network resources. It also coordinates the physical resources necessary to deliver the virtualization of the services. This is particularly visible for Infrastructure-as-a-Service. In summary, VIM is the management glue between the hardware and software resources.

B. Envisioned Video Streaming Services

The Orchestrators and Management blocks run specific methods, including the selection of the virtual services. It also defines the services chaining between the VNFs and then publishes the jobs in the queues of the concerned servers. The envisioned services are described below.

- Streaming Service: The streaming service plays a load balancer role as it receives client requests for playing specific video content and redirects the requests to the proper cache servers. The streaming server tracks also video accesses and sends statistics back to the fog nodes to be used in data analyses. This will help to improve the business intelligence of the video streaming provider.
- Transcoding Service: Responsible for generating the video stream at lower resolutions. In this way, the same video does not have to be stored in the cache with different resolutions. In a multi-tier ecosystem, a transcoding service may be deployed closer to a set of end-users and/or fog nodes, and chained with the other services if necessary.
- Cache Service: Seeks to provide the multimedia content geographically scattered over the servers. This service stores the video requested by the user as well as the transcoding service output.
- Overlay Routing Service: A set of overlay paths are built by the orchestrator or by the overlay routing service itself to avoid having a single point of failure.[12].

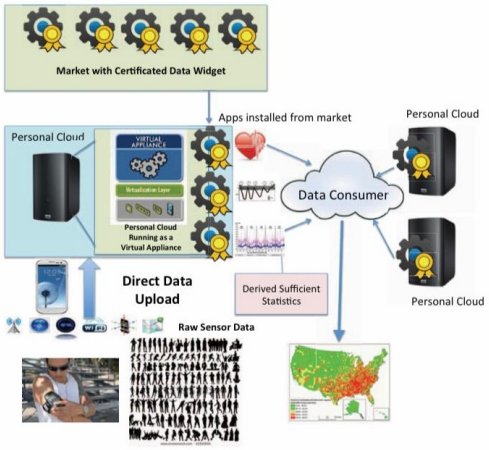
X. HYBRID CLOUD TRUST FRAMEWORK

In this framework, raw sensor data is uploaded directly to users' own personal cloud and a user has physical and

complete control of his/her own personal big data. Personal Big Data consumers such as health organizations and auto insurance companies can develop data widgets, applications to be run on users Personal Cloud, to learn behavioral patterns (e.g., average daily commuting time) from the raw sensor data, which are then uploaded to the public cloud for different applications. A third-party certificate provider certifies such a widget and inform end users what statistics it is going to generate and then upload. In this framework,users own their data. No raw data is exposed to "data consumers". It is up to users' discretion to "sell" their statistics to "data consumers" by enabling the corresponding widgets.

In this hybrid cloud approach, computation is distributed among mobile devices, personal cloud and public cloud. Raw sensor data is stored on user's personal cloud only where user has full physical control of his/her own data. We believe by doing this, users feel that the data is under control with full knowledge of where the data is stored and managed and will thus have less privacy-anxiety.

1) Private Cloud
   The contribution of the proposed framework is decentralized data aggregation. Unlike most the conventional sensing framework which has a centralized server for user data aggregation, each user uploads direct raw data to his/her private cloud. Without the user's permission, raw data will stay on the private cloud without any leakage. In our approach, private clouds are virtual machines deployed on users' own physical machines. Each virtual machine has it's built-in sensor data management system. Data Chunk: The raw data are split into data chunks before being uploaded to the private cloud. Each chunk contains a certain amount of sensing data with corresponding timestamp when collected.
2) Data Consumer and Data Widget App
   Data Consumers are third parties(trusted or untrusted) who want to leverage the statistics from user's sensor data. They cannot access users' raw data without users' permission in our proposed framework. They can develop feature extraction application, called Data Widget, which can run locally in the private cloud to extract the desired statistics and transmit the results to the cloud, only after authorized by the user. The benefit of running data widget in the private cloud is two-fold:
   - Scalability the computation is distributed to private clouds
   - Privacy the users have full control in what information to provide.
3) Trusted Third-party Data Widget Certification and App Market
   Trusted third-party data widget certification is key to the framework. There have two observations
   - Users usually do not have the knowledge to identify the features that could be reverse engineered to reveal the personal identity
   - Data consumers release widgets that behave differently from the SLD(a Data Malware). But, the third party should share no interest with all data consumers (e.g., Apple and all iOS app developers). It defines different sensitivity levels and categorizes data widgets into different levels. This trusted third-party also needs to make sure the behavior of a data widget agrees with the widget's SLD.



XI. Conclusion

This paper present survey on various cloud computing architectures and its components that provides simple, basic, and clear explanation for building a cloud environment. Cloud Computing has provided a reliable, cost-efficient, cutting-edge infrastructure for advanced software delivery models such as Software-as-a-Service (SaaS).The cloud computing middleware can help users set up virtual service environment and help enterprises transfer smoothly from traditional structure to the Cloud Computing architecture by using the cloud computing technology.CLOUDIO builds on a set of Business Information Systems scenario requirements, mostly concerned with the flexibility of the data model and managing large datasets, optimizing the underlying architecture.From a software engineering standpoint, the need for flexible, context-aware and dynamic Multi-Tenant Architectures is gaining momentum, particularly when dealing with Business Information Systems BIS) scenario requirements. Storing the data in the cloud guarantees that the data is always available and is a great way to run a bussiness.

References

[1] Song Kun, Zhoe Zhihai."The service oriented architecture[J].Marine Technology, 2007,26(4).
[2] Cloud Reference Architecture- An Oracle White Paper, November 2012, http://www.oracle.com/technetwork/topics/entarch/oracle-wp-cloud-ref-arch-1883533.pdf
[3] A. L. DeCarlo, "Cloud Service Management and Cloud Monitoring for providers: A primer".
[4] R. Iyer, R. Illikkal, Measuring, modeling and managing VM shared resources", Computer Networks, Volume 53, Issue 17, Virtualized Data Centers, December 2009, Pages 2873-2887

[5] K. Kim, A. Beloglazov, R. Buyya, "Power-aware provisioning of Cloud resources for real-time services", In Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and eScience, MGC 09, Urbana Champaign, Illinois, 2009

[6]] Linthicum D., Relevance of Enterprise Architecture to Cloud Computing, 2010./cloudsoa/2010/12/relevance-of-enterprise-architecture-to-cloud-computing.php.

[7] 2010/12/relevance-of-enterprise-architecture-to-cloud-computing.php.Cloud Security Alliance, Security Guidance for CriticalAreas of Focus in Cloud Computing, 2009.

[8] NSW Government, An introduction to enterprise architecture for records managers, 2010.

[9] Amazon Web Services, http://aws.amazon.com/

[10] Chong Federick,Carraro Gianpaolo, Wolt Roger ," Multi-Tenant Data Architecture", Microsoft Corporation.

[11] L. F. Bittencourt, M. M. Lopes, I. Petri, and O. F. Rana, "Towards virtual machine migration in fog computing," in P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2015 10th International Conference on. IEEE, 2015, pp. 1–8.

[12] R.K Sitaraman, M,.Kasbekar, W. Lichtenstein, and M.Jain,"Overlay netwoks: An akamai perspective",2014.

A Study on Web Mining Technique : Trends and Applications

K R Deepa, Sibin Babu and Sreeharsha V P
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dr Reji C Joy
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: rejic.j@vidyaacademy.ac.in)

***Abstract*—Data mining is the technique of analyzing data from different perspective and summarizing it into useful information. For extracting and analyzing the data from web, web data mining can be used . Web mining is a cross point of database, information retrieval and artificial intelligence. It can be broadly classified in to Web content mining (WCM), Web structure mining (WSM) and Web usage mining (WUM) .Data mining is work upon Off-Line whereas Web mining is work upon On-Line. The implementation of data mining technique through cloud computing will allow the users to retrieve meaningful information and thus increasing the popularity in cloud computing. The ultimate aim of this paper is to evaluate the past, present and future of web mining including the research issues, future trends, techniques and development efforts are presented in this paper.**

***Index Terms*—Keywords -Data mining, Web mining, Cloud computing, HITS, Page rank**

I. Introduction

World Wide Web(WWW) is largely popular for information's that are produced from billions of medias world wide , it's huge diverse and dynamic and thus raises the scalability, multi -media data and temporal issues. Due to those situations,currently drowning in information and facing web mining techniques could be used to solve the information overload problems directly or indirectly. However, we do not claim that Web mining techniques are the only tools to solve those problems. Other techniques and works from different research areas, such as database (DB), information retrieval (IR), natural language processing (NLP) and the Web document community, could also be used.In today's world due to the tremendous growth of information sources available on the Web and the recent interest in E-commerce.

Web mining is the application of data mining techniques to extract knowledge from Web data. The main three domains of web mining are: Web Content Mining, Web Structure Mining, and Web Usage Mining. Web content mining is the process of extracting knowledge from the content of documents or their descriptions and the process that goes beyond keyword extraction. Web structure mining is the process of inferring knowledge from the World-wide Web organization and links between references and referents in the Web. It can be subdivided into two categories based on the kind of structural data used they are hyperlinks and document Structure. Web usage mining, also known as Web Log Mining, is the process of extracting interesting patterns in web access logs. Data mining is basically discovering the unknown patterns in large amount of data. If data mining technique is applied on web it is called web mining. Therefore, Data mining is work upon Off-Line whereas Web mining is work upon On-Line. Nowadays cloud computing is gaining more popularity due to its mobility, huge availability and low cost.it also bring great security to the company's data and Information .

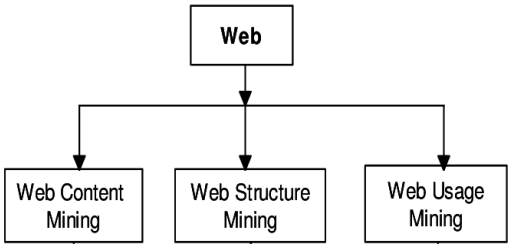


Fig. 1. Types of web mining

II. Web Mining- Overview

- Web Content mining
- Web Structure mining
- Web Usage mining

The sub tasks of web mining are the following.

- **Resource discovery** -Locates unfamiliar documents and services
- **Information selection and preprocessing** -Specific information are automatically extracted and preprocessed
- **Generalization** -Uncovering general patterns at individual Web sites and across multiple Sites.
- **Analysis** -Validation and interpretation of mined patterns.
- **Visualization** -Presenting the results of an interactive analysis in a visual

III. History of Web Mining

Web mining techniques are the result of long process of research and product development. This evolution began when business data was first stored on computers and internet, continued with improvements in data access and more recently real-time technologies that users in the WWW (World Wide Web) allow us to navigate through our data. Data gathered from surveys, or input from several independent or networked locations via Computer and tapes are define Data collection. Use the data storage typically refers to software and related activities, the retrieving, or stored in a database or other data source acting on. In the evolution from business data to business information each new step has built upon the previous one. For example, the ability to store large databases is critical to web mining. Basically data mining technique are used in web mining. Web mining is extended version of data mining. Data mining is work upon Off-Line whereas Web mining is work upon On-Line. In data mining data stored in (database) data warehouse and in web mining data stored in server database and web log.

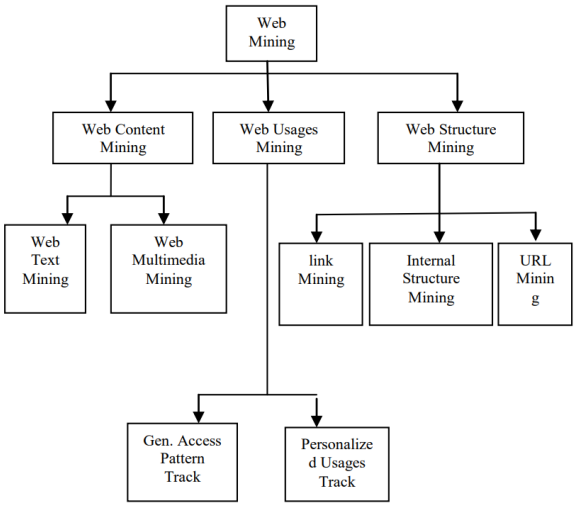


Fig. 2. Taxonomy of web mining

IV. Web Mining Categories

A. Web Content mining

Web content mining is the process of extracting knowledge content of documents or their descriptions and the process that goesbeyond keyword extraction. The Web content data consist of unstructured data such as free texts, semi-structured data such as HTML documents, and a more structured data such as data in the tables or database generated HTML pages. However, much of the Web content data is unstructured text data . The research around applying data mining techniques to unstructured text is termed knowledge discovery in texts , or text data mining , or text mining . The unstructured documents is free texts such as news stories. The bag of words or vector representation is used to represent unstructured documents. It takes a single words found in the training corpus as features. This representation ignores the sequence in which the words occur and is based on the statistic about single words in isolation. The features could be Boolean , or frequency based. Variations of the feature selection include removing the case, punctuation, infrequent words, and stop words. The features could be reduced further by applying some other feature selection techniques, such as information gain, mutual information, cross entropy, or odds ratio. Other preprocessing includes Latent Semantic Indexing that seeks to transform the original document vectors to a lower dimensional space by analyzing the correlational structure of terms in the document collection such that similar documents that do not share terms are placed in the same topic, and stemming which reduces words to their morphological roots. For example the words ”informing”, ”information”, ”informer”, and ”informed” would be stemmed to their common root ”inform” and only the latter word is used as the feature instead of the former four. While those pre-processing variations are useful for reducing feature set size, the generality of their effectiveness over different domains for text categorization tasks are doubted . Other feature representations are also possible such as using information about word positions in the document ,using n-grams representation etc. Information Retrieval View for Semi-Structured Documents The methods that are used are common data mining methods. The applications ranged from hypertext classification or categorization and clustering, learning relations between Web documents, learning extraction patterns or rules, and finding patterns in semi-structured data. Recent advances in multimedia data mining promise to widen access also to image, sound, video, etc. content of Web resources. It deals with discovering useful information or knowledge from web page contents. It is hard to make exactly the same understanding to the similar image by the different users due to multiplicity of contents in a single image and the subjectivity of human perception. To overcome this Content-Based Image Retrieval (CBIR) was invented. It helps to automate the processing of annotating image databases. The large gap between low-level image features like color histograms used to index images and high-level semantic contents of images human subjectivity is the major drawback of CBIR The user can interact with image retrieval system by two ways.

- User types in a list of keywords representing the semantic contents of the preferred images.
- The user provides a set of examples images as an input and the retrieval system will retrieve other analogous images

CBIR consist of the following steps -

- CBIR consist of a semantic network which links images to semantic annotations, that integrates semantic features and image features.
- System supports both query by keyword and query by image example through semantic network and low-level feature indexing.

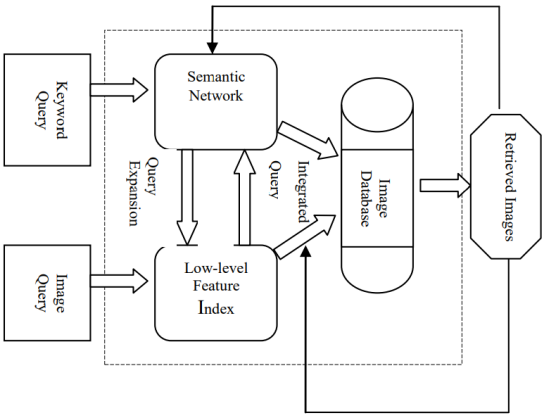


Fig. 3. Update semantic network based on user Relevance feedback

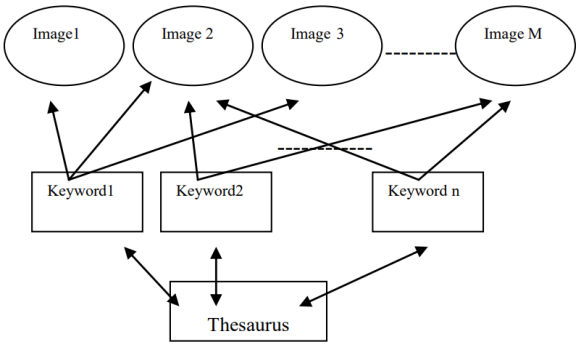


Fig. 4. Semantic network

1) Framework of integrated relevance feedback and query expansion: We can differentiate the research done in Web content mining from two different points of view: IR and DB views. The goal of Web content mining from the IR view is mainly to assist or to improve the information finding or filtering the information to the users usually based on either inferred or solicited user profiles. Web content mining from the DB view mainly tries to model the data on the Web and to integrate them so that more sophisticated queries other than the keywords based search could be performed. Basically the DB view tries to infer the structure of the Web site or to transform a Web site to become a database so that better information management and querying on the Web become possible. As mentioned previously, the DB view of Web content mining mainly tries to model the data on the Web and to integrate them so that more sophisticated queries other than the keywords based search could be performed. These could be achieved by finding the schema of Web documents, building a Web warehouse or a Web knowledge base or a virtual database. The research done in this area mainly deals with semi-structured data. Semi-structured data from database view often refers to data that has some structure but no rigid schema.

B. Web Structure mining

Web structure mining is the process of inferring knowledge from theWorld-wide Web organization and links between references andreferents in the Web. It can be subdivided into two categories based onthe kind of structural data used they are hyperlinks and documentStructure. Web structure mining tries to discover the model underlying the link structures of the Web. The model is based on the topology of the hyperlinks with or without the description of the links. This model can be used to categorize Web pages and is useful to generate information such as the similarity and relationship between different Web sites. Web structure mining could be used to discover authority sites for the subjects (authorities) and overview sites for the subjects that point to many authorities (hubs). Web structure mining focus in the structure of the hyperlinks within the Web itself, it utilizes the hyperlinks structure of the Web to apply social network analysis to model the underlying links structure of the Web itself. In research they utilizes the network analysis of people to model the network of AI researchers. They use the name entity data found in close proximity in any public Web pages such as the hyperlinks from home pages, co-authorship and citation of papers, exchange of information between individuals found in net-news archives, and organization charts. Their research could be classified as a combination of Web structure and content mining. There more applications of Web structure mining in the context of Web warehouse these include measuring the completeness of the Web sites by measuring the frequency of local links that reside in the same server and discovering the nature of the hierarchy of hyper-links in the Web sites of a particular domain to study how the flow of information affects the design of the Web sites.

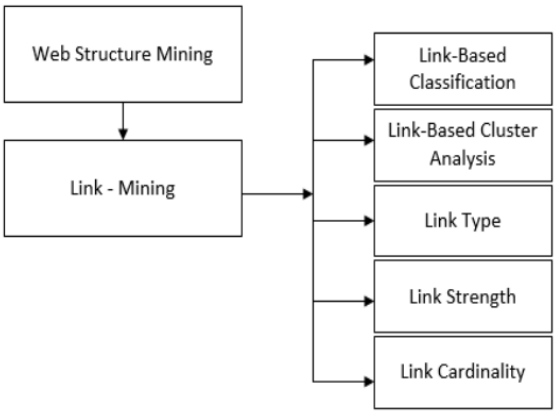


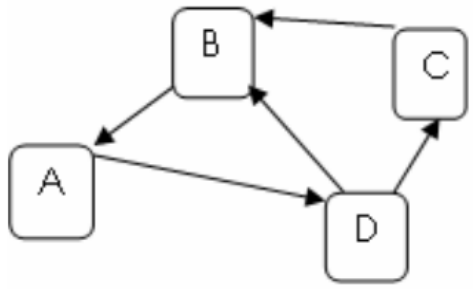
Fig. 5. Web structure mining types

- Link-based Classification: It is an upgrade classification version of classic data mining and its task is to link domains Main focus is to predict webpage categories – based on text, HTML tags, link between web pages and other attributes

- Link-based Cluster Analysis: In cluster analysis data is categorized or grouped together. Primary focus is on data segmentation
- Link Type: It helps to guess link type between entities (two or more)
- Link Strength: Link strength shows that links might be related to weights.
- Link Cardinality: Helps to find depicted website and to find comparison between them.

1) Web Structure Mining Algorithms: There are different types of web structure mining algorithms mainly two types are describing here.

- Page Rank Algorithm It was developed in 1998 by L. page and S. Brain. It is an algorithm that is frequently used to rank pages. Page rank approach leads to number of pages linking to a specific web page indicates, calculates or describes the importance of that page. It is called back link structure.



- Hyperlink Induced Topic Search. (HITS) HITS concept was developed by Jon Kleinberg. Two terminologies are used in HITS algorithm that is authorities and hubs Good authority is a page that is pointed by high hub weights and good hubs are pages that points to many authority pages with high weights.

C. Web Usage Mining

Web usage mining, also known as Web Log Mining, is the process ofextracting interesting patterns in web access logs.While the Web content and structure mining utilize the real or primary data on the Web, Web usage mining mines the secondary data derived from the interactions of the users while interacting with the Web. The Web usage data includes the data from Web server access logs, proxy server logs, browser logs, user profiles, registration data, user sessions or transactions, cookies, user queries, bookmark data, mouse clicks and scrolls and any other data as the results of interactions. Web usage mining focuses on techniques that could predict user behavior while the user interacts with the Web. As mentioned before, the mined data in this category are the secondary data on the Web as the result of interactions. These data could range very widely but generally we could classify them into the usage data that reside in the Web clients, proxy servers and servers . The Web usage mining process could be classified into two commonly used approaches . The first approach maps the usage data of the Web server into relational tables before an adapted data mining technique is performed. The second approach uses the log data directly by utilizing special pre-processing techniques. As is true for typical data mining applications, the issues of data quality and pre-processing are also very important here. The typical problem is distinguishing among unique users, server sessions, episodes, etc. in the presence of caching and proxy servers. In general, typical data mining methods could be used to mine the usage data after the data have been pre-processed to the desired form. However, modifications of the typical data mining methods are also used such as compo- site association rules , an extension of a traditional sequence discovery algorithm and hypertext probabilistic grammars .The Web usage data could also be represented with graphs . Often the Web usage mining uses some background or domain knowledge such as navigation templates, Web content, site topology, concept hierarchies, and syntactic constraints. The applications of Web usage mining could be classified into two main categories: learning a user profile or user modeling in adaptive interfaces. In practice,

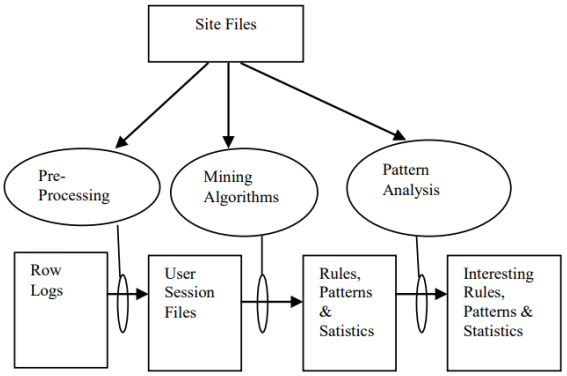


Fig. 6. High Level Web Usage Mining Process

the three Web mining tasks above could be used in isolation or combined in an application, especially in Web content and structure mining since the Web documents might also contain links. Web Usage Mining Techniques:

1) Data Preprocessing
2) Pattern Discovery
3) Pattern Analysis

1) Web Usage Mining Algorithms:

- Apriori algorithm It helps to find frequent sets of items during transaction. At first apriori algorithm observe initial database and captures those data sets which are large, then uses result of first captured data sets as a base or model to discover other data sets.
- FP Growth Algorithm FP-Growth discovers frequent sets of data from FP tree. FP tree is complete data structure, contains one root node “null” and sub tree nodes (prefix) as children. FP growth is more efficient that Apriori algorithm.

V. KEY CONCEPT

Ranking Metrics—for Page Quality and Relevance : Searching the web involves two main steps: Extracting the pages relevant to a query and ranking them according to their quality. Ranking is important as it helps the user look for "quality" pages that are relevant to the query.

A. Page Rank

PageRank is a metric for ranking hypertext documents based on their quality. The key idea is that a page has a high rank if it is pointed to by many highly ranked pages. So, the rank of a page depends upon the ranks of the pages pointing to it. This process is done iteratively until the rank of all pages are determined.

B. Hubs and Authorities

Hubs and authorities can be viewed as "fans' and "centers" in a bipartite core of a web graph, where the "fans" represent the hubs and the "centers" represent the authorities. The hub and authority scores computed for each web page indicate the extent to which the web page serves as a hub pointing to good authority pages or as an authority on a topic pointed to by good hubs. The scores are computed for a set of pages related to a topic using an iterative procedure called HITS.

C. Robot Detection and Filtering—Separating Human and Nonhuman Web Behaviour

Web robots are software programs that automatically traverse the hyperlink structure of the web to locate and retrieve information. First, e-commerce retailers are particularly concerned about the unauthorized deployment of robots for gathering business intelligence at their websites. Second, web robots tend to consume considerable network bandwidth at the expense of other users. Sessions due to web robots also make it difficult to perform click-stream analysis effectively on theweb data. Conventional techniques for detecting web robots are based on identifying the IP address and user agent of the web clients. While these techniques are applicable to many well-known robots, they are not sufficient to detect camouflaged and previously unknown robots

D. Information Scent—Applying Foraging Theory to Browsing Behaviour

Information scent is a concept that uses the snippets of information present around the links in a page as a "scent" to evaluate the quality of content of the page it points to, and the cost of accessing such a page. The key idea is to model a user at a given page as "foraging" for information, and following a link with a stronger "scent." The "scent" of a path depends on how likely it is to lead the user to relevant information, and is determined by a network flow algorithm called spreading activation.

E. Preprocessing—Making Web Data Suitable for Mining

Preprocessing of web data to make it suitable for mining was identified as one of the key issues for web mining. A significant amount of work has been done in this area for web usage data, including user identification and session creation, robot detection and filtering ,and extracting usage path patterns provides a comprehensive overview of the work in web usage data pre - processing.

F. Online Bibliometrics

With the web having become the fastest growing and most up to date source of information, the research community has found it extremely useful to have online repositories of publications. observed (Lawrence 2001) that having articles online makes them more easily accessible and hence more often cited than articles that are offline. With such information stored in the web, it becomes easier to point to the most frequent papers that are cited for a topic and also related papers that have been published earlier or later than a given paper. Fundamental web mining techniques are applied to improve the search and categorization of research papers, and citing related articles.

G. Visualization of the World Wide Web

Mining web data provides a lot of information, which can be better understood with visualization tools. This makes concepts clearer than is possible with pure textual representation. Hence, there is a need to develop tools that provide a graphical interface that aids in visualizing results of web mining. The site hierarchy is represented in a circular form called the "Disk Tree" and the evolution of the web is viewed as a "Time Tube."

VI. DATA MINING-BROAD VIEW

Data mining is the process of sorting through large data sets to identify patterns and relationships that can help solve business problems through data analysis. Data mining techniques and tools enable enterprises to predict future trends and make more-informed business decisions. Data mining is typically done by data scientists and other skilled BI (Business intelligent)and analytics professionals. But it can also be performed by data-savvy business analysts, executives and workers who function as citizen data scientists in an organization. Its core elements include machine learning and statistical analysis, along with data management tasks done to prepare data for analysis. The use of machine learning algorithms and artificial intelligence (AI) tools has automated more of the process and made it easier to mine massive data sets, such as customer databases, transaction records and log files from web servers, mobile apps and sensors. The data mining process can be broken down into these four primary stages:

- Data gathering Relevant data for an analytics application is identified and assembled. The data may be located in different source systems, a data warehouse or a data lake, an increasingly common repository in big data environments that contain a mix of structured and unstructured

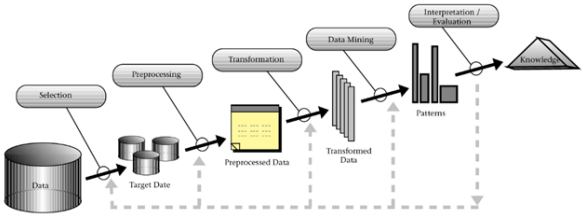


Fig. 7. Knowledge Discovery Process

data. External data sources may also be used. Wherever the data comes from, a data scientist often moves it to a data lake for the remaining steps in the process.

- Data preparation This stage includes a set of steps to get the data ready to be mined. It starts with data exploration, profiling and pre-processing, followed by data cleansing work to fix errors and other data quality issues. Data transformation is also done to make data sets consistent, unless a data scientist is looking to analyze unfiltered raw data for a particular application.
- Mining the data Once the data is prepared, a data scientist chooses the appropriate data mining technique and then implements one or more algorithms to do the mining. In machine learning applications, the algorithms typically must be trained on sample data sets to look for the information being sought before they're run against the full set of data.
- Data analysis and interpretation The data mining results are used to create analytical models that can help drive decision-making and other business actions. The data scientist or another member of a data science team also must communicate the findings to business executives and users, often through data visualization and the use of data storytelling techniques.

A. Data mining techniques

Data mining works by using various algorithms and techniques to turn large volumes of data into useful information. Here are some of the most common ones:

1) Classification
   This type of data mining technique is generally used in fetching or retrieving important and relevant information about the data and metadata. It is also even used to categories the different types of data format into different classes. If you focus on this article until it ends, you may definitely find out that Classification and clustering are similar data mining types. As clustering also categorizes or classify the data segments into the different data records known as the classes. However, unlike clustering, the data analyst would have the knowledge of different classes or clusters. Therefore in the classification analysis, you have to apply or implement the algorithms to decide in which way the new data should be categorized or classified. A classic example of classification analysis would be Outlook email. In Outlook, they use certain algorithms to characterize an email is legitimate or spam.
2) Prediction
   This technique is generally used to predict the relationship that exists between both the independent and dependent variables as well as the independent variables alone. It can also use to predict profit that can be achieved in future depending on the sale. Let us imagine that profit and sale are dependent and independent variables, respectively. Now, on the basis of what the past sales data says, we can make a profit prediction of the future using a regression curve.
3) CIustering
   In Data Mining, this technique is used to create meaningful object clusters that contain the same characteristics. Usually, most people get confused with Classification, but they won't have any issues if they properly understand how both these techniques actually work. Unlike Classification that collects the objects into predefined classes, clustering stores objects in classes that are defined by it.
4) Association rules:
   An association rule is a rule-based method for finding relationships between variables in a given dataset. These methods are frequently used for market basket analysis, allowing companies to better understand relationships between different products. Understanding consumption habits of customers enables businesses to develop better cross-selling strategies and recommendation engines.
5) Neural networks:
   Primarily leveraged for deep learning algorithms, neural networks process training data by mimicking the interconnectivity of the human brain through layers of nodes. Each node is made up of inputs, weights, a bias (or threshold), and an output. If that output value exceeds a given threshold, it "fires" or activates the node, passing data to the next layer in the network. Neural networks learn this mapping function through supervised learning, adjusting based on the loss function through the process of gradient descent. When the cost function is at or near zero, we can be confident in the model's accuracy to yield the correct answer.
6) Decision tree:
   This data mining technique uses classification or regression methods to classify or predict potential outcomes based on a set of decisions. As the name suggests, it uses a tree-like visualization to represent the potential outcomes of these decisions.
7) nearest neighbor (KNN):
   K-nearest neighbor, also known as the KNN algorithm, is a non-parametric algorithm that classifies data points based on their proximity and association to other available data. This algorithm assumes that similar data points can be found near each other. As a result, it seeks to calculate the distance between data points, usually through Euclidean distance, and then it assigns a category based on the most frequent category or average.

VII. Cloud Computing

Cloud computing is on-demand access, via the internet, to computing resources—applications, servers (physical servers and virtual servers), data storage, development tools, networking capabilities, and more—hosted at a remote data center managed by a cloud services provider (or CSP). The CSP makes these resources available for a monthly subscription fee or bills them according to usage. Organizations of every type, size, and industry are using the cloud for a wide variety of use cases, such as data backup, disaster recovery, email, virtual desktops, software development and testing, big data analytics, and customer-facing web applications. For example, healthcare companies are using the cloud to develop more personalized treatments for patients. Financial services companies are using the cloud to power real-time fraud detection and prevention. And video game makers are using the cloud to deliver online games to millions of players around the world.

- Benefits of cloud computing
  1) Agility
     The cloud gives you easy access to a broad range of technologies so that you can innovate faster and build nearly anything that you can imagine. You can quickly spin up resources as you need them–from infrastructure services, such as compute, storage, and databases, to Internet of Things, machine learning, data lakes and analytics, and much more.
  2) Elasticity
     With cloud computing, you don't have to over-provision resources up front to handle peak levels of business activity in the future. Instead, you provision the amount of resources that you actually need. You can scale these resources up or down to instantly grow and shrink capacity as your business needs change.
  3) Cost saving
     The cloud allows you to trade fixed expenses (such as data centers and physical servers) for variable expenses, and only pay for IT as you consume it. Plus, the variable expenses are much lower than what you would pay to do it yourself because of the economies of scale.
  4) Deploy globally in minutes
     With the cloud, you can expand to new geographic regions and deploy globally in minutes. For example, AWS has infrastructure all over the world, so you can deploy your application in multiple physical locations with just a few clicks. Putting applications in closer proximity to end users reduces latency and improves their experience.
- Types of Cloud computing service Models
  The three main types of cloud computing services include Infrastructure as a Service, Platform as a Service, and Software as a Service. Each type of cloud computing provides different levels of control, flexibility, and management so that you can select the right set of services for your needs.

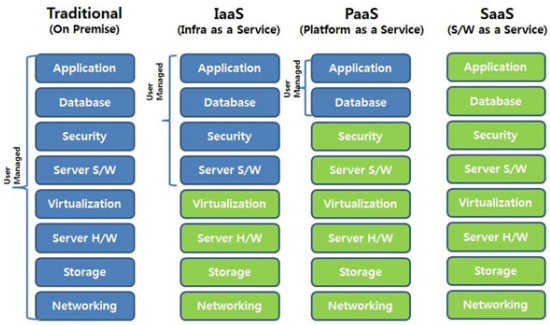


Fig. 8. Cloud Computing Service Models

  1) Infrastructure as a Service (IaaS)
     IaaS contains the basic building blocks for cloud IT. It typically provides access to networking features, computers (virtual or on dedicated hardware), and data storage space. IaaS gives you the highest level of flexibility and management control over your IT resources. It is most similar to the existing IT resources with which many IT departments and developers are familiar
  2) Platform as a Service (PaaS)
     PaaS removes the need for you to manage underlying infrastructure (usually hardware and operating systems), and allows you to focus on the deployment and management of your applications. This helps you be more efficient as you don't need to worry about resource procurement, capacity planning, software maintenance, patching, or any of the other undifferentiated heavy lifting involved in running your application.
  3) Software as a Service (SaaS)
     SaaS provides you with a complete product that is run and managed by the service provider. In most cases, people referring to SaaS are referring to end-user applications (such as web-based email). With a SaaS offering, you don't have to think about how the service is maintained or how the underlying infrastructure is managed. You only need to think about how you will use that particular software
- Types of cloud computing Deployment Models
  There are three main cloud computing deployment models: public cloud, private cloud, and hybrid cloud. Like cloud service models, each cloud deployment model has its own unique setup with a range of differing requirements and associated benefits.
  1) Public Cloud
     Public cloud is the provision of computing services over the public internet by a third-party provider. Public cloud services are available to anyone who wants to use them and can be free or paid to use.
  2) Private Cloud
     Private cloud is the provision of computing services

for a single organization access over the internet or a private internal network. A private cloud can be managed internally or by a third-party provider.

3) Hybrid Cloud
Hybrid cloud is the provision of computing services through a combination of public cloud and private cloud deployments. This type of deployment allows the sharing of data and applications between both cloud environments.

VIII. Web Mining through Cloud Computing

One of the mostly used technologies in Web Mining is Web Usage Mining. Web Usage mining using Cloud Computing is majorly adopted these days due to its reduced cost efficiency and flexibility. However, in spite of improved movement and attention, there are considerable, continual concerns about cloud computing that ultimately compromise the vision of cloud computing as a new IT procurement model. Fundamentally Cloud Mining is novel approach to faced search interface for your data. The major challenge which is a security of web mining is been offered by SaaS (Software-as-a Service) and used for dropping the cost which is termed as cloud mining technique. It's been targeted to change the existing framework of web mining to generate an influential framework by Hadoop and map Reduce communities for projecting analytics.

A. Online Web Usage Mining in Cloud System

Web based recommender system is extremely accommodating in directing the users to aim pages in particular web sites. Moreover, Web usage mining cloud model systems have been proposed to predict user's intention and their navigation behaviors. In the following, we review some of the most significant WUM systems and architecture that can be compared with our system, cloud system proposed a cloud model for navigation pattern mining through web usage mining to predict user future movements.

IX. Data Mining in Cloud Computing

Data mining techniques and applications are very much needed in the cloud computing paradigm. As cloud computing is penetrating more and more in all ranges of business and scientific computing, it becomes a great area to be focused by data mining.
Data mining in cloud computing is the process of extracting structured information from unstructured or semi-structured web data sources. The SaaS Distribution model helps to reduce costs by providing flexible license options and outsourcing the hardware effort. At SaaS, the software is not applied in the company; it lies at a software service provider's server. That means the provider deals with the hardware, looks after software updates and maintains technically everything. In Cloud Mining, the servers that provide the software are the Cloud. This can be the publiccloud from Google, Amazon.com etc, or a private cloud on the servers of a single provider. That has two main effects; on one hand the customer only pays for the tools of Data Mining he needs. That makes him save a lot compared to complex Data Mining suites that he isnot using exhaustive. And on the other hand he just pays for the costs that are generated by using the Cloud. He does not have to maintain a hardwareinfrastructure; he can apply data mining just via his browser. This reduces the barriers that keep small companies from benefiting of Data Mining

X. Application

Personalized Customer Experience in B2C E-commerce—Amazon.com
In traditional store focused on getting customer to the store thus the marketing budget is in general much higher than the instore customer experience budget. In the case of an on-line store, getting in or out requires exactly one click, and thus the main focus must be on customer experience in the store. A host of web mining techniques, such as associations between pages visited and click-path analysis are used to improve the customer's experience during a "store visit." Knowledge gained from web mining is the key intelligence behind Amazon's features such as "instant recommendations," "purchase circles," "wish-lists," etc.

A. Web Search—Google

Google14 is one of the most popular and widely used search engines. The quality and quickness of the search facility makes it the most successful search engine. Google was the first to introduce the importance of the link structure in mining information from the web. PageRank, which measures the importance of a page, is the underlying technology in all Google search products, and uses structural information of the web graph to return high quality results. The Google toolbar is another service provided by Google that seeks to make search easier and informative by providing additional features such as highlighting the query words on the returned web pages. The advertising program introduced by Google targets users by providing advertisements that are relevant to a search query. One of the latest services offered by Google is Google News16. It integrates news from the online versions of all newspapers and organizes them categorically to make it easier for users to read "the most relevant news."

B. Web-Wide Tracking—DoubleClick

Web-wide tracking is tracking an individual across all sites he visits, is an intriguing and controversial technology. It can provide an understanding of an individual's lifestyle and habits to a level that is unprecedented, which is clearly of tremendous interest to marketers. A successful example of this is DoubleClick Inc.'s DART ad management technology17. Sites that use DoubleClick's service are part of The DoubleClick Network and the browsing behavior of a user can be tracked across all sites in the network, using a cookie.

C. Understanding Auction Behavior—eBay

The eBay's22 founders was to create an infrastructure that gave this urge a global reach, with the convenience of doing it from one's home PC. it popularized auctions as a product

selling and buying mechanism. The anonymity of the web has also created a significant problem for eBay auctions, as it is impossible to distinguish real bids from fake ones. eBay is now using web mining techniques to analyse bidding behaviour to determine if a bid is fraudulent .

D. Personalized Portal for the Web—My Yahoo

Yahoo23 was the first to introduce the concept of a "personalized portal," i.e. a web site designed to have the look-and-feel and content personalized to the needs of an individual end-user. Mining My Yahoo usage logs provides Yahoo valuable insight into an individual's web usage habits, enabling Yahoo to provide personalized content, which in turn has led to the tremendous popularity of the Yahoo web site.

XI. Conclusion

Web mining is viewed based on three categories web content mining, web usage mining, web structure mining and also it provides the information of web research mining Algorithms . Web data mining is growing constantly and so much research is carried out in that area. The paper is evaluated on the current web mining research availableto analyze web data and extract all manner of useful knowledge from it.We have briefly described the key computer science contributions made by the field, a number of prominent applicationsand outlined some areas of future research. Web mining is the extension of data mining . Data mining has a great scope in the future in the field of web service optimization, fraud and threat analysis, identification of web matrices etc. More research is carried out in this field based on the requirements of web application user.The other new trend is the use of cloud computing.Cloud Computing Internet services will rely on clouds of servers to handle tasks. Data mining in cloud computing is the process of extracting structured information from unstructured or semi-structured web data sources. Here we explore the how the data mining tools like SAS, PAS and IaaS are used in cloud computing to extract the information. Leading cloud computing providers Amazon Web Services, Windows Azure, OpenStack. Whenever we work on mining over cloud computing that time we hesitate for the cost but that come very less by cloud mining. So, we can say that cloud mining can be seen as future of web mining.

References

[1] Short Survey of Data Mining and Web Mining using Cloud Computing Int. J. Advanced Networking and Applications Volume: 12 Issue: 05 Pages: 4725-4731(2021) ISSN: 0975-0290
[2] Performance Analysis of Web Mining Techniques and Future Directions IJCST Vol. 3, Issue 2, April - June 2012
[3] Web Mining Research: A Survey volume : 2, Issue : 1, Pages : 15
[4] Web Mining Concepts, Applications, and Research Directions, Jaideep Srivastava, Prasanna Desikan, Vipin Kumar
[5] Data Mining: Web Data Mining Technique, Tools and Algorithm : An Overview volume : 9, No.6,2018
[6] S. Chakrabarti, Mining the Web: Discovering Knowledge from Hypertext Data (Else-vier Science & Technology Books, 2003), ISBN-13: 9781558607545.
[7] R. Chau, C. Yeh and K. Smith, Personalized multilingual web content mining, KES(2004)

E-commerce Security Issues and Solutions

Aysha P A, Dona James and Krishnapriya T S
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dijesh P
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: dijesh.p@vidyaacademy.ac.in)

***Abstract*—The emergence of e-commerce has changed the way business is done world-wide. Opportunities have arisen for initiating new international and domestic business ventures. As individuals and businesses increase information sharing, vulnerability to attack or intrusion rises. Therefore, security is a necessity in an E-commerce transaction. The purpose of this paper is to investigate and present the new approach to E-commerce security. Our philosophy is to look for significantly new paradigms and shifts from previous thinking, and facilitates the debate within a constructive environment of new approaches to existing and on going global problems of tackling security issues in E-commerce. There is a remarkable scope for more streamlined living through an increase of e-platforms especially ecommerce, but this coincides with an increase in security concerns since the global market place is virtual and anonymous. Therefore, users have to trust the online providers blindly. To overcome this physiological barrier, the e-platforms should ensure utmost security. If not the e-commerce, industry is unable to perform in the market effectively. Thereby arises the need to carry out a systematic review of security issues in the e-commerce industry and to discover how different frameworks address these problems.With the emergence of the Economy, and with an ever-increasing percentage of consumers doing their business primarily via e-commerce, is fast being regarded as the way to go global at the touch of a button. Hence, developing an effective E-Commerce model is becoming vital for any modern business. E-Commerce Security has some main issues. They are interception of data, redirection of data, identification of parties, exploitable program errors, and being the weakest point in security. When administrating a secure e-commerce site, it is important to remember that you are part of a link of systems.This paper presents the conceptual components of the e-commerce in general, and identifies and classifies the different types of security challenges and risksin e-commerce.**

***Index Terms*—E-commerce,SSL,Authenticity,Digital Signature**

I. Introduction

E-commerce means Electronic Commerce that refers to any kind of transactions that is conducted online. The companies and individuals can sell their things through internet. As its conducted through internet there are a lot of challenges in e commerce. One of the main challenge is Security. The transactions are done in different locations using a personal computer or workstations. E commerce is also done through wireless devices such as mobile devices. The wireless devices helps users to communicate, buy or sell things from anywhere at anytime. E-commerce can be drawn on many technologies such as mobile commerce, Internet marketing, online transaction processing, electronic funds transfer, supply chain management, electronic data interchange (EDI), inventory management systems, and automated data collection systems.

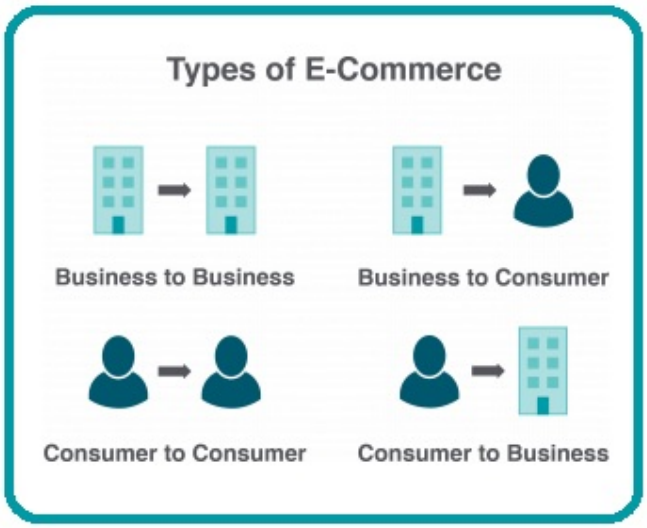


Fig. 1. Types of E-commerce

E-commerce has become a critical component in any business competitive strategy. Organisations are gaining opportunities and benefits such as global presence and improved competitiveness from Web based e-commerce. However, exploring these opportunities challenges conventional notions of business management, because e-commerce changes the characteristics and rules of business competition through electronic flows of information and money.

II. Overview of E-commerce

E-Commerce refers to the activity of buying and selling things over the internet. Simply, it refers to the commercial transactions which are conducted online.E-commerce has emerged as a solution to the traditional business that has only advantages and no disadvantages. But it has some disadvantages. The main vulnerable points in e commerce include the

hardware, software and the environment. Hardware security refers to the devices that run e commerce websites. They must provide proper firewall securities. So before implementing e-commerce, we have to check the security of our systems and the data.

III. Types of E-commerce

A. Business to Consumer(B2C)

The term business-to-consumer (B2C) refers to the process of selling products and services directly between a business and consumers who are the end-users of its products or services. Advantages of B2C:

- Elimination of intermediaries – Businesses, particularly manufacturers, can offer lower and more affordable goods to consumers by selling products directly, eliminating distributors and retailers that add to the cost of the product.
- Flexibility in pricing – Price tags can be adjusted easily and instantly, benefiting the business and the customer. You also have the ability to strategically cross-sell, discount, and provide coupons and other online/offline specials.
- Professional image – Even if you have a small business, your e-commerce site can enhance your reputation by projecting a larger picture and allowing you to compete on a level playing field.
- Extended reach – Opening up an online store front can effectively extend your presence to a great number of prospects, particularly those unable to access your local bricks and mortar store.

B. Business to Business(B2B)

B2B E-commerce, or business-to-business electronic commerce, is the sale of goods or services through online transactions between businesses. Because orders are processed digitally, buying efficiency and effectiveness is improved for wholesalers, manufacturers, distributors and other types of B2B sellers. Advantages of B2B:

- Purchasing supplies – By creating an online account for your business with supply stores you can purchase office supplies and equipment online and save time and money by automating the purchasing process.
- Purchasing direct materials – These are materials that going to the production or manufacturing of your products. Establishing a relationship with a vendor and purchasing online may help reduce costs.
- Selling products or services to new vendors – Having an online e-commerce presence opens up more opportunities to extend your reach and gather new vendors beyond your bricks and mortar offerings.

IV. Digital E-commerce Cycle

Security is very important in online shopping sites. Now days, a huge amount is being purchased on the internet, because it‘s easier and more convenient. Almost anything can be bought such as music, toys clothing, cars, food and even porn. Even though some of these purchases are illegal we will be focusing on all the item‘s you can buy legally on the internet. Some of the popular websites are eBay, iTunes, Amazon, HMV, Mercantile, dell, Best Buy and much more. In this figure the user enter data in the computer if all information is correct then next procedure is done then next user enter the credit and debit card information for payment the Order . Then next order is complete the email is sent to customer and merchant and next is company send the order of product to customer home.

V. Security Issues

Now a days, lot of people are engaged in online shopping, almost everything they are bought from online. So as the reason it consist of many websites are involved, such as eBay, iTunes, Amazon, HMV, Mercantile, dell, Best Buy and much more. So as digital E-commerce consist of a cycle which has different phases such as Order placed by user, Shopping cart, Payment, Order is completed, Email sent to customer, Sent to warehouse for fulfilment, Shipping to customer.

With the fast growth of the internet, it has become the platform for the criminal attacks and intrusions. And these attacks spread very fast and they can come from anywhere in the world. As we know that when the term E-commerce is defined the very first perspective that comes to the mind is security. The internet is a platform that consist of many private networks connected together. These private network is then exposed to the public network which has a large scope for threats. Protection against these threats should have a very strong security

There are following types of security issues in any commerce application which needs to be addressed

- Viruses: They have ability to replicate and spread to other files.
- Worms: They are designed to spread from computer to computer.
- Trojan horse: They appears to be benign, but then does something other than expected.
- Bots: It can be covertly installed on computer; responds to external commands sent by the attacker.
- Adware: It calls for unwanted pop-up ads.
- Spyware: It can be used to obtain information, such as a user’s keystrokes, e-mail, IMs, etc.
- Phishing and Identity Theft: Any deceptive, online attempt by a third party to obtain confidential information for financial.

VI. Risk In E-commerce Security

There are several risk involved due to inadequate security such as Bugs or miss-configuration problems in the web server that can cause the theft of confidential documents risks on the Browsers’ side i.e. breach of user’s privacy, damage of user’s system, crash the browser etc, Interception of data sent from browser to server or vice versa. The psychological aspect of security incorporates the feeling of fear, the need to feel that

one's money is secure, and the ability to control the payment process and performance of online transactions.

WAP is the protocol about security. It tells that Wireless Transport Security Layer(WTLS) provides a secure infrastructure for m-commerce applications. During an HTTP request, wireless requests to web pages are translated at WAP gateway from WTLS protocol to SSL protocol. Encryption and decryption is done during translation. Some of the addressing for software security are:

- Security risks of wireless devices must be carefully analysed and addressed.
- "WAP gap"
  - Wireless requests to web pages are translated at the WAP gateway from the WTLS protocol to SSL protocol, widely used in HTTP requests.
  - If an attacker compromises the WAP gateway, could capture data when decryption is done.
- WAP gap problem
  - Should be solved by simple modifications to existing protocols.
- Low level languages
  - In handheld devices cause the continuation of basic flaws like Buffer overflow etc.
- Application developers may forgo security features like encryption etc
  - Due to Limited power, lack of Processing cycles, memory and bandwidth of the devices
  - To increase online performance.
- Interesting software development
  - The ability to send & execute mobile code.
- WML script is used to overcome software application risks.

VII. Security Factors

Security is an essential part of any transaction that takes place over the internet. Customer will lose his/her faith in e-business if its security is compromised. Following are the essential requirements for safe e-payments/transactions

- Confidential : Information should not be accessible to unauthorized person. It should not be intercepted during transmission.
- Integrity : Information should not be altered during its transmission over the network.
- Availability : Information should be available wherever and whenever requirement within time limit specified.
- Authenticity : There should be a mechanism to authenticate user before giving him or her access to required information.
- Non-Repudiabilty : It is protection against denial of order or denial of payment. Once a sender sends a message, the sender should not able to deny sending the message. Similarly the recipient of message should not be able to deny receipt.
- Encryption : Information should be encrypted and decrypted only by authorized user.
- Auditability :Data should be recorded in such a way that it can be audited for integrity requirements.

VIII. Measures to Ensure Security

- Digital Signature :Digital signature ensures the authenticity of the information. A digital signature is a e-signature authenticated through encryption and password. It helps in E- commerce.
- Security Certificates : Security certificate is unique digital id used to verify identity of an individual website or user. Due to this we can know about the user details.
- Secure Socket Layer(SSL) :SSL(Secure Sockets Layer) is the standard security technology for establishing an encrypted link between a web server and a browser. This link ensures that all data passed between the web server and browsers remain private and integral.
- Secure Hypertext Transfer Protocol (SHTTP) : SHTTP extends the HTTP internet protocol with public key encryption, authentication and digital signature over the internet.

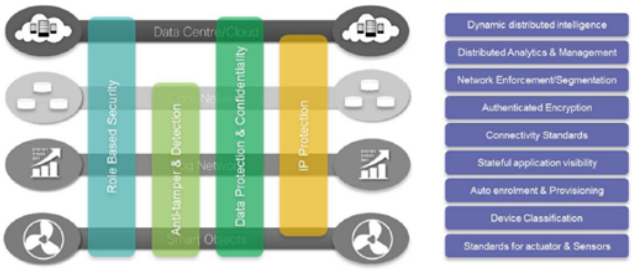


Fig. 2. CISCO Security Framework

As the trend of online transactions continues to grow, there is an increase in different types of attacks against security. E-commerce security is the protection of e-commerce assets from unauthorized access, use, modification or destruction. The existence of systems weakness may be exploited by an attack could cause the platform to enter an unsafe state. The vulnerability points are customer end, server side, and communication channel. Lack of security may hinder people from using e-commerce. Therefore, security must achieve both robustness and fault tolerance by preventing the attacks. There are different dimensions of security like privacy, authentication, integrity, non-repudiation and availability. The main security issues faced by both customers and providers are transactional security, privacy, system security and cyber crime. This paper provides a review of 4 different frameworks about security.

- CISCO Security framework
  In case of transactional security, the Information should not be stolen or revealed. In case of privacy, no personal data should be used without owner's consent. There

are database threats in system security. Cyber crime includes identity theft, fraudulent activities, virus attacks. CISCO frameworks provides anti tamper, detection, data protection and security based access.

- Constrained Application Protocol Framework (CoAP)
  CoAP consists of modules to handle security and trust issues that works in application layers. It provides packet protection, duplicate detection and availability.
- Floodgate Security Framework

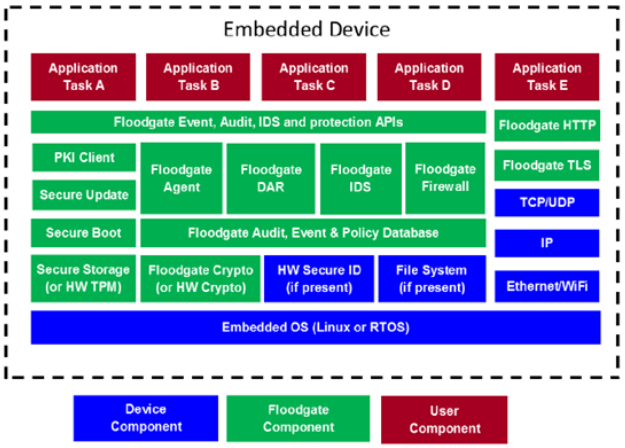


Fig. 3. Floodgate Security Framework

This helps in securing cyber threats by building secure, authenticated and trusted services. In case of the security issues, the framework provides some modules like IDS module, PKI Client and Firewall module.

- Object Security Architecture Framework
  It provides security within the application payload. The framework provides authentication, confidentiality, and privacy through capability based access control.

IX. Guidelines for Healthy Online Shopping

If a customer wants to use e-commerce website then following things need to be checked.

- A URL address that begins with https://inplaceofhttp://
- By Symbol (lock icon) or by a message
- A Browser that notify the customer that he/she is in secure site
- Facts which minimizing security and privacy risks in e/m-commerce
- By reviewing security features provided by ISP, web design, software company
- Pay attention towards security alerts and install security patches
- By update, scan for spyware and viruses
- By taking backup system and information
- Obtain a digital certificate for website
- Avoiding storing transaction information/account information in system
- By developing privacy policy
- By verify customer addresses
- By fraud filters to screen for suspicious activity

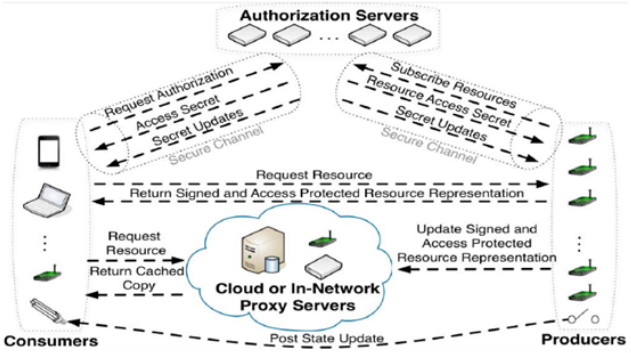


Fig. 4. Object Security Architecture Framework

X. Privacy in E-commerce

Privacy is an interdisciplinary issue. The right of humans for keeping their privacy is debated in many fields, including the areas of law, politics, philosophy, sociology, and more recently computer sciences. This may either be necessary for the e business transaction itself. Whenever a customer sends a message he should ensure that they are communicating through correct server, what they send is delivered un-modified, they can prove that they sent the messages. All of the concerns listed can be resolved using some combination of cryptographic method, and certificates methods.

There are different policies used to ensure and measure security in E-commerce environment

- Privacy polices architecture the manner in which a company collects, uses, protects data, and the choices they offer consumers to exercise rights when their personal information is used. The basis of this policy, consumers can determine whether and to what extent they wish to make information available to companies
- Cryptography Secret key cryptography is the oldest type of method in which to write things in secret. There are two main type of Cipher systems are classified into 2 classes which are:- Secrete key cryptography, transposition and substitution. Transposition cipher, encrypt the original message by changing characters order in which they occurred, where as in substitution cipher, the original message was encrypted by replacing the characters with other characters. In both types, both the sender and receiver share the same secret keys. The most widely used secret key scheme today is called Data Encryption Standard (DES). DES cipher work with 56-bit secret key and 16 rounds to transform a block of plaintext into cipher text. Public key method use two different but mathematically related keys. One of the keys is used to encrypt the data, i.e. plaintext and the second key is used to decrypt the cipher text .

- Certificates are the next concept in ecommerce it binds identity, authority, public key, and the other information to a user. For most internet E commerce application, certificates using a format defined in international telecommunication union telecommunication standardization sector.
- Pretty Good Privacy (PGP) provides a confidentiality and authentication service that can be used for electronic mail and file storage applications. PGP has grown explosively and is now widely used, three main reasons can be cited for this growth. First: It is based on algorithm that has survived extensive public review and are considered extremely secure. Second: It has a wide range of applicability. The actual operation of PGP consists of five services: authentication, confidentiality, compression, e-mail compatibility and segmentation.

XI. Conclusion

Security is very important if you are to succeed online. Hackers are getting better at their games, which means you need a dedicated team that will stay updated with security issues and provides around-the-clock protection to your websites.In e-banking security system SSL and SET (Secure Socket Layer and Safely electronic transaction) communication protocols are used. SSL and SET solve the safety problem in electronic payment of credit card. They ensure the confidentially of information. Reduce deployment costs and distribute information easily. The Future scope of the study of Security is use to reduce threats. By using Security issues wages and salary bill of banks get reduced Online banking convenience of client has considerably increased as the can transact from home or office. If the better security operate then net banking and e- marketing will be increase.

References

[1] Kuldeep Kaur Int. Journal of Engineering Research and Applications www.ijera.comISSN : 2248-9622, Vol. 5, Issue 5, (Part -6) May 2015, pp.63-73.

[10] Rui Wang, Shuo Chen "How to Shop for Free Online Security Analysis of Cashier-as-a-Service Based Web Stores". IEEE S&P'11 proceedings.

[2] The Study of E-Commerce Security Issues and Solutions Sangeetha M K Prof. Dr. Suchitra RMS(IT),IV Sem Head of the dept Jain University, Bangalore Jain University, Bangalore.

[3] Corbitt, B.J. ," Trust and e-commerce a study of e-commerce perceptions", Electronic Commerce Research & Application, Vol .2 No.3 ,pp.203 -15(2003).

[4] Analysing risk of security and its challenges in e commerce by DivyaSree.

[5] A Systematic Review of Security in Electronic Commerce Threats and Frameworks By N. Kuruwitaarachchi, P.K.W. Abeygunawardena, L.Rupasingha & S.W.I.Udara.

[6] Haya Alshehri, FaridMeziane, "The Influence of Advanced and Secure E-Commerce Environments on Customers Behavior: The Case of Saudis in the UK," in 12th International Conference for Internet Technology and Secured Transactions, 2017.

[7] ShaziaYasin, Khalid Haseeb. "Cryptography Based E-Commerce Security: A Review". IJCSI-Vol. 9, Issue 2, No 1, March 2012.

[8] V.SRIKANTH "ECOMMERCE ONLINE SECURITY AND TRUST MARKS". IJCET ISSN 0976 – 6375, Volume 3, Issue 2, July- September (2012).

[9] Seyyed Mohammad Reza Farshchi "Study of Security Issues on Traditional and New Generation of E-commerce Model" International Conference on Software and Computer Applications-IPCSITvol.9(2011).

[11] Haya Alshehri, Farid Meziane, "The Influence of Advanced and Secure E-Commerce Environments on Customers Behaviour: The Case of Saudis in the UK," in 12th International Conference for Internet Technology and Secured Transactions, 2017.

[12] Cong Cao, Jun Yan, Mengxiang Li, "The Effects of Consumer Perceived Different Service of Trusted Third Party on Trust Intention: An Empirical Study in Australia," in 14th IEEE International Conference on e-Business Engineering, 2017.

[13] Sheshadri Chatterjee, "Security and Privacy Issues in E-Commerce: A Proposed Guidelines to Mitigate the Risk," in IEEE International Advance Computing Conference, 2015.

[14] Kima S ," An Empirical Study of Customers perceptions of security and trust in e-payment system," Electronic Commerce research and applications, Vol .9 No. 1 pp.84-95(2010)

[15] Joseph P.T, S.J(2008)," An Indian perspective", 3 rd edition, E-Commerce,by PHI learning private limited.

[16] Aladwani.A," Key Internet Characteristics and e-commerce issues In Arab Countries", Information Technology and People ,Vol 16,1:9-20(2003)

Health Care, Security Standards and Applications in IoT

Arjun K, Lishell Rapheal and Merin V Wilson
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Sajay K.R
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: sajay.k.r@vidyaacademy.ac.in)

***Abstract*—Constructing smart systems based on IOT combined with cloud computing is an emerging trend. Today an enormous number of researches are going to apply IOT in different areas of our life. One of the major IOT applications is health care. IoT enabled health care systems aim in better treatment quality, and reducing its cost for the users. Now a days health care uses a queuing system which takes time to give treatment to patients in hospital. To process such data efficiently, we can use Multicore technology. It describes two or more CPUs working together on the same chip. It supports parallel computing. Parallel computing improves performance by reducing the time of processing. During the long waiting time, many sensor stream data can lose due to the fixed size of buffer. This will reduce waiting time for treatment. According to IoT technology has brought about revolutionary impacts in many areas of our lives. IoT-based smart technologies which allow individuals to connect and control smart devices and appliances remotely using computers, smartphones, or tablets through the internet. Connecting one's devices and appliances to the Internet however, exposes them as well as the data sensed, collected, and exchanged by them to a wide range of security threats and risks. opportunities, applications and challenges of IoT. New opportunities in IoT will meet business requirements, and new services will be created based on real-time physical world data. IoT bring new changes to our lives and it can be applied in many fields like industries, agriculture, health care etc. IoT includes challenges like technical challenges, architecture challenges, privacy and security challenge.**

***Index Terms*—IOT, DAS, UUID, RFID, SixLoWPAN, Multi-core, HIPAA, FERPA, PCI-DSS, CMMC, C2M2, ad hoc manner.**

I. Introduction

The Internet of Things (IoT) is regarded as a technology and economic wave in the global information industry after the Internet. The IoT is an intelligent network which connects all things to the Internet for the purpose of exchanging information and communicating through the information sensing devices in accordance with agreed protocols. It achieves the goal of intelligent identifying, locating, tracking, monitoring, and managing things. It is an extension and expansion of Internet-based network. IOT expands the communication from human and human to human and things or things and things. In the IoT paradigm, many objects surrounding us will be connected into networks in one form or another. RF identification (RFID), sensor technology, and other smart technologies will be embedded into a variety of applications.

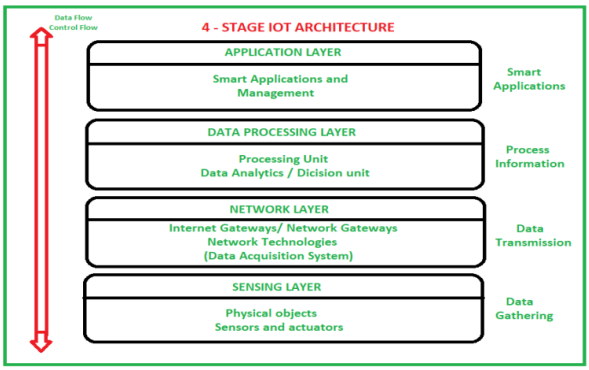


Fig. 1. ARCHITECTURE OF IOT

II. Architecture of IoT

The architecture of IoT depends upon its functionality and implementation in different sectors. Still, there is a basic process flow based on which IoT is built. The fundamental architecture of IoT has 4 Stage IoT architecture.

A. Sensing Layer

Sensors, actuators, devices are present in this Sensing layer. These Sensors or Actuators accepts data, processes data and emits data over network.

B. Network Layer

Internet/Network gateways, Data Acquisition System (DAS) are present in this layer. DAS performs data aggregation and conversion function. Advanced gateways which mainly opens up connection between Sensor networks and Internet.

C. Data processing Layer

This is processing unit of IoT ecosystem. Here data is analysed and pre-processed before sending it to data centre from where data is accessed by software applications often termed as business applications where data is monitored and managed and further actions are also prepared.

D. Application Layers

This is last layer of IoT architecture. Data centre or cloud is management stage of data where data is managed and is used by end-user applications.

III. Emerging Trends and Parallel Processing of IoT Applications

A. IoT Enabling Technologies

1) Sensors: Sensors and actuators are able to sense the environment automatically, gather the data and transmit data to specialized applications for the process and bring it to action.

2) UUID (a universal unique identifier): In IoT system each object has a unique digital identity so that their tracking can be done easily. This technique of assigning every object with a unique id is known as a universal unique identifier (UUID).

3) RFID (Radio-frequency identification): A typical RFID system includes an RFID tag and an RFID reader. Because of its capability of identifying, tracing, and tracking, the RFID system has been widely applied in different health care systems. It could send thing's real-time information in IoT, which can be found extremely helpful for extending doctors services to the patients.

4) Cloud computing: Continuous collection of data through tracking and sensing leads to a very large amount of data that has to be processed, stored and shared so comes the need of cloud computing.

B. Applications

1) Child Care : IoT can be used extensively in monitoring the total growth, vital parameters and behaviour of children. These systems can alert medical personnel in real time about under-weight cases or abnormal readings of vital parameters in order to take further medical diagnosis.

2) Emergency Care: During emergency conditions, all what is needed is the information about the facilities and the hospitals based on particular treatment given.

3) Disaster management through smart ambulance: There can be significant increase in probability of saving their lives if the needed vital signs of patients are transferred in real time to the hospital while they are still on board the ambulance.

C. SixLoWPAN Protocol

SixLoWPAN stands for IPv6 over Low power Wireless Personal area networks. It originates from IoT to make smallest low power devices with limited processing capacity to interact with internet. This protocol is similar to IEEE8202.15.4 to communicate in wireless medium. IPv6 is not supported in IEEE802.15.4 protocol. SixLoWPAN supports IPv6. Since IoT uses IPV6 as its addressing scheme to allocate a unique ID to each and every device. This protocol provides mobility model in health care applications. It provides exchange of messages between patient nodes, base stations and visited networks. It is used for active communication.

D. Parallel Processing of Data Packets

IoT data packets are generated in rapid speed. Normally the data packets are buffered in a fixed sized queue. Data lost occur if the device process data packets slowly. To reduce the overhead to the processing device, the queue size is limited. Data is streamed in such a manner to process packets in a proficient manner. Stream level partitioning is done on an existing system to do complex event processing. In the existing methodology Complex Event handling was done to utilize the processing device efficiently to reduce the traffic flow in IoT and to process time critical packets in an efficient manner. However, it is not sufficient for processing medical data. Delay in processing is handled with the help of Multicore technology. Multicore technology is used to parallelize the execution. It splits the execution of the main task into sub tasks and multiple cores executes each sub task parallel. This technology can be used to process medical data in an efficient manner. It speeds up the processing and reduces the execution time.

IV. A Review of Security Standards and Frameworks for IoT-based Smart Environments and Security Testing Methodology of IoT

A. Health Insurance Portability and Accountability Act (HIPAA)

HIPAA was developed to provide guidelines for enabling health plans, health care providers and health care clearing houses to implement sufficient controls for securing employee or customer health information and protect sensitive patient health information from being disclosed without the patient's consent or knowledge.

B. Family Educational Rights and Privacy Act (FERPA)

FERPA was developed to protect the privacy of student education records and applies to all schools that receive funds under an applicable program of the U.S. Department of Education

C. Payment Card Industry Data Security Standards (PCI-DSS)

PCI DSS was designed to help protect the safety of card data and defines a set of requirements intended to ensure that all organisations that process, store, or transmit credit card information maintain a secure environment to reduce credit card fraud.

D. Cyber security Maturity Model Certification (CMMC)

CMMC is used to measure defence contractors' capabilities, readiness, and sophistication in cybersecurity.

E. Cyber Security Capability Maturity Model (C2M2)

Developed by the U.S. Department of Energy (DOE) C2M2 enables organizations to voluntarily measure the maturity levels of their cybersecurity capabilities consistently.

F. NIST Cyber Security Framework

The NIST cybersecurity framework was created based on a set of industry standards and best practices to help organizations manage their critical infrastructure cybersecurity risks.

G. NIST Risk Management Framework (RMF)

The Risk Management Framework (RMF) provides a comprehensive, flexible, repeatable, and measurable 7-step process (prepare, categorize, select, implement, assess, authorize, and monitor) that any organization can use to manage information security and privacy risks.

H. NIST Privacy Framework

The NIST privacy framework was developed to help organizations identify and manage privacy risks as well as build innovative products and services while protecting individuals' privacy.

I. Testing Methodology

For testing the integrity of the IoT network one need to follow a testing methodology. To keep the security principles of IoT intact, testing of the IoT network and the device is very important. This testing methodology makes the IoT network resilient to external attacks. The below figure shows the testing methodology of IoT Network:

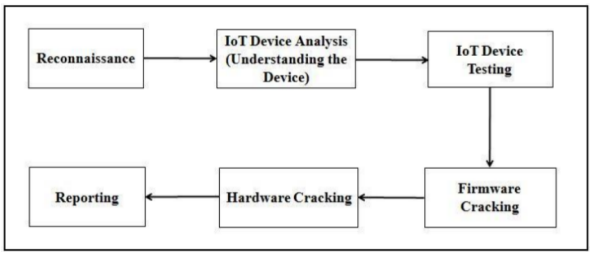


Fig. 2. Testing methodology of IoT Network

1) Reconnaissance : Reconnaissance is a phase in which information gathering is done. It is an important phase as it helps us narrow down our target IoT devices or IoT network.

2) IoT Device Analysis (Understanding the Device): Understanding the working of the IoT device means understanding the behaviour of the device. The working also includes the functional working of the device. This helps in envisaging the security threats that could prevail in the IoT device.

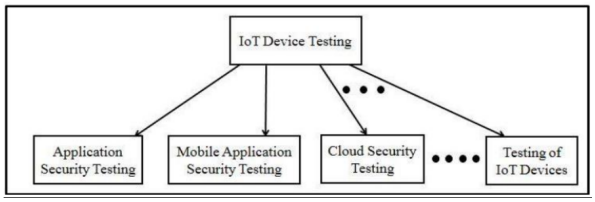


Fig. 3. IOT Device Testing

3) IOT Device Testing :

- Application Security Testing
  The application or device has web interface then the application security testing would be performed.
- Mobile Security Testing
  The application or device has mobile application for communication then one would perform Mobile Application Security Testing
- Cloud Security Testing
  The data which is exchanged between the two devices in the IoT network, store data on the cloud. Therefore, in cases like these, Cloud Security Testing is also performed.

4) Firmware Cracking : This cracking is crucial because the firmware is badly packed and coded. There are times where the important credentials and other port details are stored in cleartext. This crucial data can be of great use to the attacker. The firmware is easy source of data sometimes as they are easily available over the internet. They are also sometimes provided with the IoT device. One should not discard firmware cracking while testing IoT.

5) Hardware Cracking : Hardware cracking involves opening of the device. The memory chips are removed from the devices and are read using EPROM readers and other chip readers. The chips store sensitive information in them which are read by the attacker. Hardware cracking is sometimes not reversible. Once dismantled, cannot be assembled again and brought back to the working state.

V. IoT Opportunities, Applications and Challenges

A. Opportunities

The IoT has a vast of opportunities. In IoT, applications, services, middleware components, networks, and end nodes will be structurally organized and used in entire new ways. New opportunities will meet business requirements, and new services will be created based on real-time physical world data.

B. Applications

1) Smart Agriculture: Agricultural resources utilization, production and cultivation of environmental monitoring, safety and traceability of agricultural product.

2) Smart Logistics: Inventory control, distribution management, traceability and smart logistics.

3) Smart Grid: Monitoring of power facilities, smart power, smart scheduling and remote meter reading.

4) Smart Environmental Protection: Pollution source monitoring, water quality monitoring and air quality monitoring.

5) Smart Medical: Medicine control, hospital management, collection and analysis of medicine parameters.

6) Connected Cars: A connected car is a vehicle which is able to optimize its own operation, maintenance as well as comfort of passengers using onboard sensors and internet connectivity.

7) Smart Retail: Retailers have started adopting IoT solutions and using IoT embedded systems. It improves store operations such as increasing purchases, reducing theft, enabling inventory management, and enhancing the consumer's shopping experience.

8) Smart Supply Chain: Supply chains have already been getting smarter now a days. It Offer solutions to problems like tracking of goods, helping suppliers, exchange inventory information.

C. Challenges

1) Architecture Challenge: Communications among the devices in IoT are expected to happen anytime, anywhere. So, it should be in a wireless, autonomic, and ad hoc manner.

2) Technical Challenge: IoT technology can be complex for variety of reasons. First, there are heterogeneous architectures in the existing network. Second, communication technologies should be of low cost and with reliable connectivity.

3) Hardware Challenge: hardware researches should be designed with low size, low cost yet sufficient functionality.

4) Privacy and Security Challenge: security and privacy issues of IoT become more important. It includes privacy of users, devices, networks etc.

5) Standard Challenge: Standards play an important role in forming IoT. A standard is essential to allow all actors to equally access and use.

D. Security Challenges

1) Encryption Capabilities: Data encryption and decryption is a continuous process. The IoT network's sensors still lack the capability to process. The brute force attempts can be prevented by firewalls and segregating the devices into separate networks.

2) Privacy Issues: IoT is all about the exchange of data among various platforms, devices, and consumers. The smart devices gather data for a number of reasons, like, improving efficiency and experience, decision making, providing better service, etc.; thus, the end point of data shall be completely secured and safeguarded.

3) Common Framework: There is an absence of a common framework and so all the manufacturers have to manage the security and retain the privacy on their own. Once a common standardized framework is implemented, the individual efforts will then collectively be utilized in an expandable manner and so reusability of code can be achieved.

4) Automation: Eventually, enterprises will have to deal with more and more number of IoT devices. This enormous amount ofuser data can be difficult to manage. The fact cannot be denied that it requires a single error or trespassing a single algorithm to bring down the entire infrastructure of the data.

5) Updations: Managing the update of millions of devices needs to be adhered to, respectively. Not all the devices support over the air update and hence it requires manually updating the devices. One will need to keep a track of the available updates and apply the same to all the varied devices.

VI. Conclusion

This paper discusses different means of tracking patient information through Iot enabled devices and applications, by which the necessary healthcare can be given to a patient through child care, emergency care etc. This paper is an attempt to categorize the applications of IoT in health care., SixLoWPAN technology, parallel processing of data packets.security standards and assessment frameworks that can be deployed in IoT-based smart environments are quite different from those that can be used in non IoT domains, the need for effective security standards and assessment frameworks for IoT-based smart environments is now inevitable. This is backed up by the fact that IoT-based smart environment security is dependent on a wide range of security checks which many existing security standards and assessment frameworks discussed in this study. The methodology discussed in the paper is suitable for any IoT device scenario. Along with the testing methodology, firmware cracking is also showcased which is part of the testing methodology. It shows the gravity of security assessment to be performed as sensitive information like Telnet cleartext usernames and passwords are revealed for the device. This paper emphasis on opportunities' challenges, and applications of Iot. The future of IoT will be expected to be unified, seamless, and vast. IoT have become an inevitable trend of development of information industry. The opportunities and applications of IOT are spread over a vast area. lives. There are many security threats and requirements that need to be dispatched.

References

[1] Nazia Ahmad,Remya P George,Rifat Jahan, "Emerging Trends in IoT for Categorized Health care", International Conference on Intelligent Computing,Page number:978, May, 2019.

[2] K.Niranjana Devi,R.Muthuselvi ,"Parallel Processing of IoT Health care Applications",October, 2013.

[3] Abhishek R.Chandan,Dr.Vaishali D.Khairnar," Security Testing Methodology of IoT" ,International Conference on Inventive Research in Computing Applications,Page number:1431, March,2018.

[4] Nickson M.Karie,Nor Masri Sahri,Wencheng Yang,Craig Valli,Victor R.Kebande," Review of Security Standards and Frameworks for IoT Based Smart Environments", Digital Object Identifier, vol. number:9,Page number:121975,September, 2021.

[5] Shanzhi Chen,Hukcheng wang,"Vision of IoT:Application,Challenges and Opportunities", Internet of Things, vol. number:4,Page number:349, August, 2014.

[6] Mohit Kumar Saini,Rakesh Kumar Saini, "Internet of Things:Applications ans Security Challenges", International journal of Engineering Research and Technlogy, vol. number:7, March, 2019.

[7] Abdur Rahim,"IoT and Cloud Convergence:Opportunities and Challenges",IEEE World Forum on Internet of Things,Page number:375, November, 2013.

[8] Hoa Tran Dang,Dong Seong Kim,"An Information Framework for Internet of Things: Services in Physical Internet",Page number:43967, August, 2018.

[9] S Narasimha Swamy,Solomon Raju Kota,"An Empirical Study on System Level Aspects of Internet of Things",Digital Object Identifier,Page number:188082, October, 2020.

[10] Omar Said,Zafer Al Makhadmeh,"Energy Management Scheme for Green IoT Environments", vol. number:8,Page number:44983, February, 2020. D. V. Dimitrov. (2016, Jul). "Medical Internet of Things and Big Data

[11] in Healthcare." Health Inform Res. [Online]. 22(3), pp. 156-163. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4981575/ [Jul. 31, 2016].

[12] J. A. Stankovic, "Research directions for the Internet of Things," IEEE Internet Things J., vol. 1, no. 1, pp. 3–9, Feb. 2014.

[13] Ruben Mayer, Boris Koldehofe, and Kurt Rothermel," Predictable LowLatency Event Detection with Parallel Complex Event Processing", IEEE INTERNET OF THINGS JOURNAL, VOL. 2, NO. 4, AUGUST 2015.

[14] G. Cugola and A. Margara, "Processing flows of information: From data stream to complex event processing," ACM Comput. Surv., vol. 44, no. 3, pp. 15:1–15:62, Jun. 2012.

[15] G. Kortuem, F. Kawsar, D. Fitton and V. Sundramoorthy "Smart objects as building blocks for the Internet of Things" IEEE Internet Comput., vol. 14, pp. 44-51, 2010 .

[16] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," Computer Networks, vol. 57, 2266-2279, 2013.

[17] P. N. Mahalle, B. Anggorojati, N. R. Prasad, and R. Prasad, "Identity authentication and capability based access control (iacac) for the internet of things," J. of Cyber Security and Mobility, vol. 1, 309-348, 2013.

[18] S. Pinto, J. Cabral, and T. Gomes, "We-care: An IoT-based health care system for elderly people," Proc. IEEE Int. Conf. Ind. Technol., pp. 1378–1383, 2017.

[19] D. Ruiz-Fernandez, D. Marcos-Jorquera, V. Gilart-Iglesias, V. Vives- ´ Boix, and J. Ram´ırez-Navarro, "Empowerment of patients with hypertension through BPM, IoT and remote sensing," Sensors (Switzerland), vol. 17, no. 10, 2017.

[20] S. Misbahuddin, J. A. Zubairi, A. R. Alahdal, and M. A. Malik, "IoT-Based Ambulatory Vital Signs Data Transfer System," J. Comput. Networks Commun., vol. 2018, pp. 1–8, 2018.

Big Data: Analysis and Applications

Adarsh K Sunny, Frederic Sojan and Rohini T B
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Manesh D
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: manesh.d@vidyaacademy.ac.in)

***Abstract*—Big data analytics helps businesses to get insights from today's huge data resources. People, organizations, and machines now produce massive amounts of data. Social media, cloud applications, and machine sensor data are just some examples. Big data can be examined to see big data trends, opportunities, and risks, using big data analytics tools. This paper gives an overview of the role of big data analytics in the field of supply chain management , logistics and transportation industry, Airline and also we discusses some solution and services that helps to improve the quality and efficiency of the application that uses big data analytics. The paper also explain about the services and design issues used in smart city project and discusses the privacy and security challenges in big data analytics.**

***Index Terms*—Big Data Analytics, Supply chain Management, Logistics and Transportation, Smart City , Real-Time Analytics, Airlines, Security and Privacy**

I. Introduction

Big data analytics is the use of advanced analytic techniques against very large, diverse data sets that include structured, semi-structured and unstructured data, from different sources, and in different sizes from terabytes to zettabytes. Big data is a term applied to data sets whose size or type is beyond the ability of traditional relational databases to capture, manage and process the data with low latency. Big data has one or more of the following characteristics: high volume, high velocity or high variety. Artificial intelligence (AI), mobile, social and the Internet of Things (IoT) are driving data complexity through new forms and sources of data. For example, big data comes from sensors, devices, video/audio, networks, log files, transactional applications, web, and social media — much of it generated in real time and at a very large scale.

In this paper, we propose to give a review of the latest applications of big data analytics in the field of logistics and transportation industry. Now a days , there are many challenges in logistics and transportation industries due to the emergence of E-commerce. They are using a huge data source to generate a huge quantity of unstructured data daily, to deal with such complex data, the use of big data analytic tools becomes an obligation. This paper is about the latest applications of big data analytics in the field of logistics and transportation industry and it also mentioning Hadoop big data analytics system and also propose a novel approach to detect and recognize containers code based on a Hadoop big data analytics system.

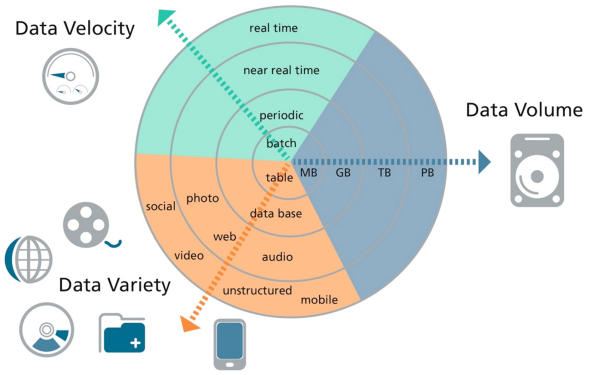


Fig. 1. 3 V's Big Data

Technological revolution in the recent past has enabled the concept of Smart City for urban development. Smart City concept is conceived with the objectives of providing better services to the citizens and improves the quality of life. This paper gives an overview about the role of big data in building a smart city. Real time big data analytics helps in making better decisions and accurate predictions. The paper also explains some of the solutions and services that uses big data analytics in developing smart city. Analytics represent the ability to make data-driven decisions, based on a summary of relevant, trusted data, often using visualization in the form of graphs, charts and other means. Supply chains typically generate massive amounts of data. Many companies are planning to extend the existing IT-Infrastructure with Big Data solutions, due to the challenges they face like the ongoing globalization, increasing data volumes and customer needs in products and services. New suppliers and distributors have to be integrated into the supply chain. Additionally, the supply chain operations and processes have to be constructed for global application by considering new product variants. The focus of each company lays more in their context and the supply chain environment. Both are not only influenced by the market itself, but also by risks or bottlenecks that might occur along the supply chains or

within production lines. Big Data solutions have to be adapted to the company's requirements and environment.

In this study, we aim to model efficiency evaluation by a deeper understanding of each process in the planning and execution phases, which a specific function in an airline represents each of these processes and finally the paper ends by discussing about the security and privacy challenges in big data analytics. The paper organizes as, Section 2, brief about overall review about big data. Section 3, discusses about how big data analytics used in supply chain management. Section 4, explain about the categorization and potential benefits. Section 5, brief about data analytics techniques used in logistics and transportation industry. Section 6, explains the services and techniques used to build a smart city. Section 7, discusses the importance of using big data analytics in Airlines and Section 8, summarize the privacy and security challenges in big data analytics.

II. Literature Review on Big Data Analytics

Big data analytics is used in many areas such as machine learning, computer vision, web statistics, medical applications, DNA analysis, data classification and clustering and in public and private industry including the logistics and transportation industry.

A broad search was conducted to document the number of papers published in the area of 'Big data analytics' the search period used was from 2012 to 14 March 2018. An overview of the number of articles is presented in Fig.2 After looking in to the journal's specific search, Fig.3 clearly demonstrates that the Annals of Operations Research was the leading journal in terms of the numbers of papers from Scopus.

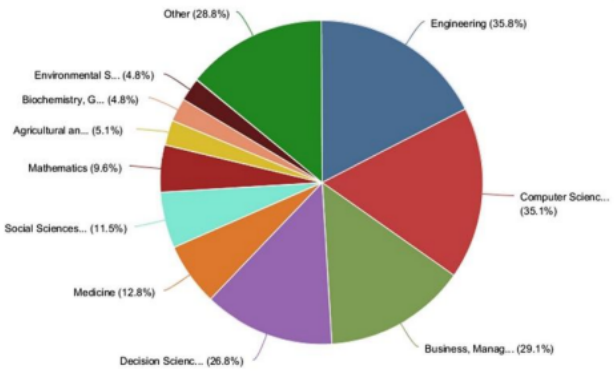


Fig. 2. Subject classification of published big data analytics papers

To refine the search to identify the document type, the Scopus databases were investigated further. In this search, conference papers, notebook chapters, books, and short surveys were excluded. After excluding these items, a total of 313 were considered. The document types considered in this section included articles, articles inpress, and reviews. The number of big data analytics papers contributed according to country of origin was also analyzed; the top ten countries are,United States led the list with 94 documents, followed by the China with 88 documents, United Kingdom with 36 documents, India with fifteen documents, and Germany with fourteen documents.The increasing interest in this research area provides solid evidence that this SI is timely and is especially relevant for the authors and readers of theTransportation Research Part E: Logistics and Transportation Review. In addition, the subject classification of published big data analytics papers is shown in below figure: 2

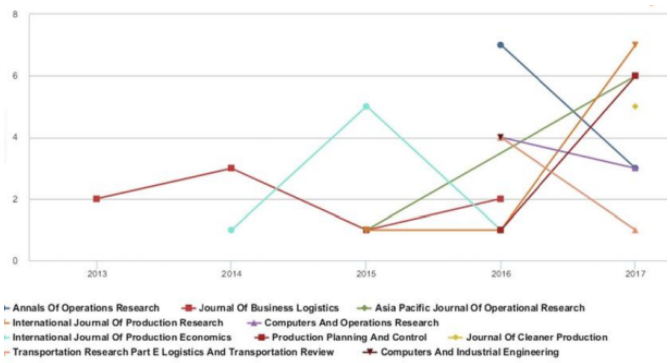


Fig. 3. Annals of Operations Research

III. Big Data for Supply Chain Management

Big data describes a way of collecting, managing and analysing large amounts of data. So, it is known by three V's, ie,volume, velocity and variety.The volume characterizes the vast amounts of data stored within the IT infrastructure. Velocity describes the large amounts of data that arrive in real-time irregularly. variety relates to the various structure of information handled within the big data environment. One of the core developments of the SCM is the SCV(Supply Chain Visibility).The SCV is like a complex problem happens due to the interaction between involved people, processes, technologies and information flows. The main target of the SCV is to show current activities and involvements along a supply chain (Eg: Collection of information for decision makers).Due to the arising outdated data sets issues, some big data solutions for supply chain management are given. LOD and SW used to increase SCV. And for describing the SCM ,the business transactions are evaluated and analysed by collecting the external data, those results are updated to the supply chain network information.This collected information is used to know upcoming risks and interruptions within the supply chain.In supply chain the companies exchange their data,but they also have an individual semantic understanding of the given information.

IV. Categorization and Potential Benefits

A Classification

There are seven potential scenarios regarding the classification of the analysed type of data delivers, and these

scenarios use both, structured and unstructured data, and can be further categorized in click stream-, social media-, server-log-, sensor-, location-, text-, and video-/audio-data. Big Data Analytics applications can contribute to different targets and provide benefits in various fields.These sections involves operational efficiency, customer experience and new business models. Operational efficiency based on Big Data capabilities uses data for better decision-making and improvement in terms of process quality, performance or resource consumption. There are RFID technologies and bar codes etc which helps in many read-points to be shared along the supply chain.

B Potential Benefits

A Capgemini study [22] identifies machine-tomachine-communication (M2M) as the issue with the most gained importance in the last year. M2M data is derived from sensors attached on machines or products, which can collect various types of data like temperature, usage data or positioning data. M2M enables automatic information between different objects like, cameras, transport vehicles, containers, etc . The use case of automatic ordering can be fully automated up to a selfdistribution of necessary goods. Regarding to the above mentioned classification, machine-to-machine communication will enable new business models and has the potential to highly increase operational efficiency. smaRTI project aims to increase the intelligence of the material flow, so the identification, localization of handling units. Like palettes, is achieved through AUTO-ID technologies, like, barcode or RFID reader which leads to better transparency and sped up processes in real-time among companies.

V. Big Data Analytics for Logistics and Transportation Industry

There are many big data analytics project for logistics and transportation to deal with the huge data coming from vehicle sensors, GPS devices etc. Some of these projects are mentioned here.

- Real-time vehicles monitoring in India : In this project, the authors collect data from vehicles about fuel, speed, acceleration, GPS location coordinates using vehicles sensors and GPS devices with other data such date, time, driver's id, etc. and then send these information by packets over wireless communication (GPRS) to clustered servers running Hadoop. All this unstructured data comes from hundreds of vehicles sending packets every 2 seconds to a HDFS system to store such big data. An analysis is made then weekly or monthly over these terabytes of data using Hadoop analytics system in order to improve the transportation company productivity and help reduce the costs.
- City of Dublin, Public Transit System : IBM big data analytics helped the city of Dublin to improve its public bus transportation network and reduce the increasing traffic congestion problem. The main idea was to reduce congestion in the city of Dublin in its public bus transport network without making major modification for the city historic infrastructure. The project helped the city to better monitor and manage the traffic in real time, which accelerate decision making and improved traffic flow and mobility in the city
- City of Da Nang, Vietnam, Traffic Management System : This project is made by IBM smart city technologies in order to reduce traffic congestion and pollution in city of Da Nang in Vietnam . The objective of the project is to build water and transportation traffic management system able to deal with the city fast-growing population of a city of more than a million population. All the different information that are collected is combined together in big data analytics system that provides a real-time summary of traffic status, events and incidents through maps and alerts for better management of the city transportation network.
- British Airways' Know Me Program : The objective is to understand clients' needs better than any competitive airline company. The project collected different types of information about 20 million customers via many sources and after that the big data analytics system stores and analysis all the structured and unstructured data to identify customers' preferences, characteristics, problems, and to provide them high quality services.
- City of Stockholm Real-Time Intelligent Transportation Services : In this project they used big data analytics system combined with the collected data with past traffic data and weather forecasts to generate more accurate predictions about future traffic conditions such as shortest-time routes in real-time.
- Cloud-Enhanced System Architecture for Logistics Tracking Services : The goal of this project is to build a system that combine internet of things, SaaS cloud architecture and big data analytics technologies for setting-up an efficient real-time monitoring of customers cargoes.

VI. Real Time Big Data Analytics in Smart City And It's Design Issues

- Smart transportation and parking: It was mainly for providing safe and quality public transport system, better traffic management and well-maintained parking facility. Real-time analytics can be used in traffic management system which is integrated with geographical information system enabled city road map.
- Smart environment and health: Early detection of natural disasters like flood, earthquake is much essential to avoid the huge loss in the city. So these applications must use real-time analytics on related data generated for timely prediction of the possibility of such natural disasters. Real-time big data analytics in healthcare results in improving the quality of healthcare monitoring system. Realtime analytics can provide a better report on the status of patients.
- Smart security and safety: A smart city project must provide better solutions for public security issues like

preventing crimes, cyber attacks, terrorist attack etc. It is important to predict the possibility of crime in the city at right time for providing the better public safety in the city.

- Smart infrastructure and agriculture: Well established infrastructure in the city is the basic need for the development of city. Smart grid and energy for power distribution and management, intelligent building and complexes, effective waste management, efficient road network and parking system are collectively formed smart infrastructure of the city. A smart city project should also focus on agriculture by identifying agriculture location within the city and adjacent to the city, which in turn help in reducing the living cost of the city.

There are four important stages to design a smart city ,they are :

1) Data collection and pre-processing : Most smart city project use the data available in social media applications, devices like sensors, video surveillance cameras, some smart applications designed for the public service. . Data pre-processing is to be done on collected data to improve the data quality. Many data pre-processing techniques like data cleaning, integration, reduction, transformation are being used appropriately at this stage.
2) Data analysis and Data Science: In this stage, different data science technologies are to be applied to prepared data for further analysis. The common methods used in many cases are clustering and classification model.
3) Prediction and data driven decisions: Once the data streams are analysed with appropriate data analysis models, it can be used to find out desired predictions and decisions using predictive models.
4) Application integration and services: Integrating data driven decisions and predictions to the end user applications designed for the city service.

VII. Proposed Big Data System for Containers Code Recognition

Ship transportation is very active industry with very large number of containers to be transported every day. Unique identifiers codes are written on the container to supervise the delivery of these containers, but manual reading of these codes include lot of problems such as slow speed and high error rate. An automatic container code recognition require three steps: text detection, characters extraction and finally text recognition. To overcome the high computation time of the employed text detection method and text recognition method, we decided to use Hadoop MapReduce to have a parallel execution model as a solution to reduce the computation time for the proposed system. First, container code is captured using monitoring cameras or mobile devices and stored on Hadoop distributed system file (HDFS). After that a pre-processing and graying color image steps are applied on the captured image. Next, we decompose the gray image in 20x20 pixel blocs. These blocs are analyzed and classified separately on different machines using MapReduce programming model to extract the text regions. Next, another step is used to separate code characters from the extracted text regions. Then, Optical Character Recognition (OCR) is applied on individual characters using MapReduce programming model. Finally, container code is recognized by merging these characters(Refer figure : 4).

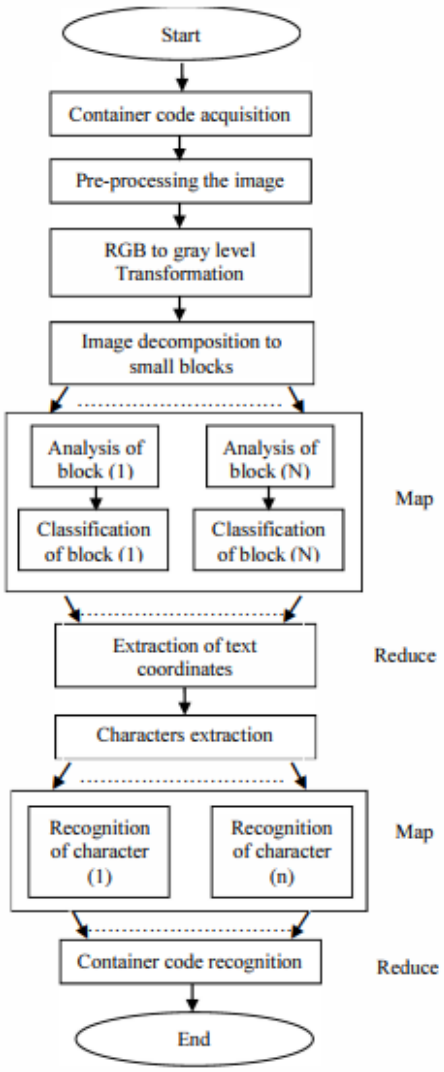


Fig. 4. Big Data System for Container recognition

VIII. Big Data Analytics in Airlines

The airline industry is well-known as a high-risk industry with a slight profit. Airline schedule planning is a crucial issue in airline operation typically designed most efficiently and effectively. Irregularities due to weather, aircraft defects, natural disasters can cause additional schedule revision increasing the deviation between operational planning and execution; these events will result in higher operational cost and lower operational efficiency. To minimize the cost of

irregularities, researchers have formed two approaches, recovery approaches and robust planning approaches. Many studies develop algorithms to optimize flight planning, rostering optimization, and for flight recovery optimization in D-day operation. Those studies aim to ensure that airline operation meets the most efficient plan and execution. We synthesize articles related to airlines, scheduling process, Multi-Criteria Decision Making (MCDM), data envelopment analysis, and data analytics. Various methods have been proposed to assess airlines in term of performance evaluation, efficiency evaluation, or benchmarking. The models usually have many criteria for decision making, thus, called the Multi-Criteria Decision Making (MCDM) model. There are two major categories:

1) The model which has many criteria and weights based on expert justification or historical data.
2) The model with criteria and mathematical programming approach and Data Envelopment Analysis (DEA).

Many researchers have discussed Multi-Criteria Decision Making (MCDM) using weighted criteria justification approaches applied in the airline industry. Many articles discuss efficiency in airlines measured using Data Envelopment Analysis (DEA), which firstly introduced by Charnes. We synthesize the extant studies related to flight schedule planning and recovery; there are several parameters used by researchers for measuring airline efficiency and performance. These parameters are: The number of Aircraft and Crew,The no of flights, Block time, Connection time etc. The data are collected from various databases of diverse enterprise applications like; Schedule Planning System (SPS), Crew Management System (CMS), Aircraft Communications Addressing and Reporting System (ACARS), Flight Planning System (FPS),etc. In order to data analytics, we perform a series of data processing, including data extraction, data cleaning, modeling, calculation, and visualization to get meaningful operational business insights. Data extraction process consists of Creation of database connection, The query for flight movement data crew data, Creates database view for flight details. The data cleansing process ensures that data taken from databases are free of error or anomaly such as incorrect movement timing, negative duration of flight time or air time, null value, duplicate data while validation ensures that the data used is relevant within the specified period when the object is serviceable. In the data modeling process, we construct object models like flight, crew, and cost. Next, the model includes related parameters which can be used to monitor operational performance indicators or further calculations.

IX. Big Data Analytics: Security and Privacy Challenges

Big Data is characterized by any storage that has large volumes, high velocity, valid, varied and valuable content such as navigation history, cookies, social networks, camera records, watched videos and gathered personal information. The processing of such data is made possible by using multiple techniques, called Big Data Analytics, which allow getting enormous benefits by dealing with any massive volume of unstructured, structured and semi-structured content that is fast changing and impossible to process using conventional database techniques. Big Data Analytics is the use of advanced analytic and parallel techniques to process very large and diverse records that include different types of contents. However, while Big Data represents an immense opportunity for many industries and decisions makers, it also represents a big challenge for respecting privacy and security concerns. This challenge arises from the fact that these analytics tools consist of storing, managing, analysing, visualizing, and sharing varied data gathered from all possible and available sources. Privacy and security of these collected records are of very importance and high priority. Big data is a new concept that describes any volume of heterogeneous data that is not possible to process by using the conventional database techniques. It consists of different kinds of digital content: Structured Data, Semi-Structured Data, Unstructured Data. A good big data is measured by some criterias like Volume, variety, veracity, value, velocity. Different Analytic techniques are used in Big Data Analytics.Apache Hadoop is the first and the most known tool for processing big data. Hadoop analytic tool is an open source software mainly conceived for big data. It is a non-relational database architecture dealing with a very big amount of distributed heterogeneous data. Some other well-known techniques are MapReduce, HDFS, Hive, Cassandra etc. There are some set of Privacy and Security concerns that must be considered before building a big data environment. They are:

1) Random Distribution: The main challenge with big data solutions is to be able to distribute storage and processing according to the regulations and data sensibility.
2) Privacy: Current big data analytics treat all data with the same priority and do not associate special actions, like encryptions or blind processing , to that kind of data.
3) Computations: It is important to secure and protect these computations to avoid any risk or attempt to change or skew the extracted results. It is also important to protect the systems from any attempt to spy on the nature or the number of performed computations.
4) Integrity: Before searching for insights and making decisions based on big data, it is important to ensure the validity and the trust level of that data in order to avoid relying on a suspect or compromised records.
5) Communication: Big data is stored in several nodes belonging to many clusters which are distributed around the world. All communications between clusters and nodes are assured through ordinary public and private networks.
6) Access Control: In a big data context, access to the data should be managed by a strong access control system to deny any malicious party from getting access to the storage servers. That is, the only node with sufficient administrative rights could have the possibility to manage and process any content.

X. Conclusion

The availability of Big Data, low-cost commodity hardware, and new information management and analytic software have

produced a unique moment in the history of data analysis. The convergence of these trends means that we have the capabilities required to analyse astonishing data sets quickly and cost-effectively for the first time in history. These capabilities are neither theoretical nor trivial. They represent a genuine leap forward and a clear opportunity to realize enormous gains in terms of efficiency, productivity, revenue, and profitability.As more and more data is generated and collected, data analysis requires scalable, flexible, and high performing tools to provide insights in a timely fashion. However, organizations are facing a growing big data ecosystem where new tools emerge and become outdated very quickly. Therefore, it can be very difficult to keep pace and choose the right tools. In this paper, we highlight a review based on some studies on big data analytics, then explains about the importance of big data analytics in supply chain management followed by its categorization and potential benefits. Moreover, we discussed about the services and solutions used in the logistics and transportation industries and seen some of examples of smart city projects and its design issues. Finally, we explained how big data analytics used for efficiency evaluation using DEA in Airlines and also discussed the privacy and security challenges in big data analytics.

REFERENCES

[1] Abdelkarim Ben Ayed, Mohamed Ben Halima, Adel M. Alimi, "Big Data Analytics for Logistics and Transportation ," 4th IEEE International Conference on Advanced logistics and Transport (ICAl T), 311-316, 2015.

[2] Manjunatha,Annappa B,"Real Time Big Data Analytics in Smart City Applications,"International Conference on Communication, Computing and Internet of Things (IC3IoT),279-284,2018.

[3] Jenes Leveling,Matthias Edelbrock,Boris Otto,"Big Data Analytics for Supply Chain Management,"Proceedings of 2014 IEEE IEEM,918-922,2014.

[4] Youssef Gahi,Mouhcine Guennoun,Hussein T Mouftah,"Big Data Analytics:Security and Privacy Challenges,"IEEE Symposium on Computer and Communication(ISCC),2016.

[5] Zudha Aulia Rachman,Arviansysh,"Big Data Analytics in Airline:Efficiency Evaluation using DEA,"7th International Conference on Information and Communication Technology(ICoICT),2019.

[6] R. Lea,"Smart cities: An overview of the technology trends driving smart cities,"IEEE Advancing Technology for Humanity,March 2017.

[7] N. Marz and J. Warren,"Big Data : Principles and best practices of scalable realtime data systems," Manning publications,2013.

[8] Big Data Analytics Options on AWS, Available: https://docs.aws.amazon.com

[9] https://www.ibm.com/in-en/analytics/hadoop/big-data-analytics

[10] https://www.techtarget.comsearchbusinessanalytics/definition/

A Literature Review of Big Data

Mariya C Johny, Saranya V Simon, and Silpa Sasi
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Siji K B
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
email : siji.k.b@vidyaacademy.ac.in

***Abstract*—This paper starts with the changes of the current social sectors in the Big Data era, and outlined the concept of Big Data. This paper focuses on the processing technology and processing tools about Big Data, mainly elaborate from four aspects, including the data preprocessing, data storage and management, data analysis and privacy protection. Finally, on the basis of the previous analysis, the trends of Big Data processing technology are given, which provides a reference for the research of Big Data processing technology.Here provide an overview of a few of the issues that can arise when analyzing big data: the inability of many off-the-shelf packages to scale to large problems; the paramount importance of avoiding suboptimal access patterns as the bulk of processing moves down the storage hierarchy; and replication of data for storage and efficiency in distributed processing. The challenges posed by data volume is most visible. Scientists catch the computing limitation constantly due to the varying data volume. Web applications produce huge amount of data because they want to manage or they want to deal with lots of customers.Top five security and privacy challenges are discussed. Each issue is attached with its own use cases.Achieving data security in big data has, therefore become one of the most important barriers that could slow down the spread of technology. Without adequate security guarantees, big data will not achieve the complete level of trust. There are four different aspects of Big Data security: infrastructure security, data privacy, data management, and integrity and reactive security.**

***Index Terms*—Big Data, Characteristics, Security, Tools ,Techniques, Privacy, Challenges, Pathologies, Implementation, 5V's**

I. Introduction

In recent years, with the development of information technologies such as the Internet, wireless transmission, network communications and smart wearable devices, a large amount of data has been collected in almost every aspect of our lives and the global data volume has dramatically increased exponentially. Nowadays the technologies of Big Data has been applied to many industries, such as government administration, network communications, healthcare, financial markets, weather forecasting, traffic management and news reporting.

Big Data is defined as "a mass, high-speed, data-rich information asset that requires efficient, new forms of information processing to enhance insight, make good decisions and optimize processes". Big Data can also be defined as information assets, pointing out that Big Data refers to data set that can't be captured, managed and processed by conventional software tools within a certain time range. It defines as massive, high growth and diversified information assets that requires new processing models to have better decision-making power, insight, and process optimization capabilities.

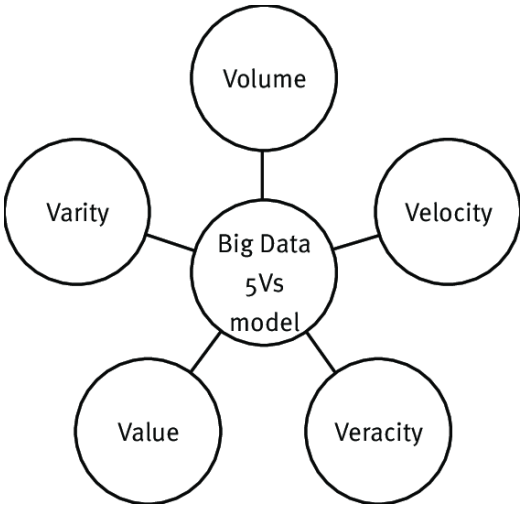


Fig. 1. The 5V model of Big Data

Some of characteristics can be seen from the definition of Big Data, such as large-scale data sets, different structure types, need fast processing speed, low value density but great value information, etc. The describe of Big Data characteristics has been formed from 5V

1) Volume Means the large-scale data sets. This is the most important features of Big Data. The data produced nowadays is estimated in the order of zettabytes, and it is growing around 40 percentage every year.
2) Value Big Data's value density is low but value information is great. We can dig out some hidden information and create more value from Big Data by analysis and processing. But what's really useful in analytics is just a part of it, not all data is available, resulting in a low value-density.
3) Variety Indicates the various types of data, which includes semi-structured and unstructured data such as audio,

video, webpage, and text, as well as traditional structured data.

4) Velocity Means the timeliness of Big Data, specifically, data collection and analysis must be rapidly and timely conducted, so as to maximize the use of the commercial value of Big Data.
5) Veracity There is the need to check the accuracy of the data by eliminating the noise through methodologies such as data pedigree and sanitization. This is to ensure data quality so that decisions that are made from the collected data are accurate and effective.

II. Big Data Processing: Techniques and Tools ,Data Processing

According to the process of Big Data processing, Big Data processing technologies involves four aspects: data preprocessing, Big Data storage and management, Big Data mining and Big Data privacy protection. There are many techniques in Big Data preprocessing, which can be summarized into four contents, data cleaning, data integration, data reduction and data transformation,

A. Data cleaning

mainly to solve the 203 inconsistency in the data, such as by filling in the empty data, identifying or removing outliers to clear the noise in the data.

B. Data integration

is to combine the data stored in multiple sources into a single data source, data integration helps to reduce the redundancy of the data set, such as data warehouse.

C. Data reduction

techniques reduce the size of the data set and ensure the integrity of the data by reducing the data. Therefore, it is more effective to mine the reduced data set.

D. Data transformation

is to meet the needs of the analysis, the data conversion under certain conditions, does not affect the results of data processing.

III. Big Data Techniques

A. Storage and Management

Due to the large data size, complex types and the high level services, Big Data brings a big challenge to the storage system. File system and database technology are commonly used to store and manage data. Distributed file systems can effectively store and manage large-scale data files in an extensible manner. In the field of distributed file system research, Google started earlier and developed Google File System (GFS) [17] in 2003. Other companies and re-searchers have also developed their own file storage systems based on different storage needs

1) Apache Hadoop-HDFS:

2) FaceBook-Haystack:

3) Ali Cloud-Pangu File System:

4) NoSQL database: is becoming a standard for Big Data because of its support for massive data, semi-structured and unstructured features. There are three mainstream NoSQL databases [22]: Key-value Storage Databases, Column Storage Databases, and Document Storage Databases at present.
Key-value storage is a simple data storage model, the data is stored in the form of key-value pairs. Column storage database is used to store and process data in a column storage structure, which is mainly suitable for batch data processing and real-time query.Document databases can support a much more complex data structures than key-value storage, and are common with SimpleDB, MongDB, and CouchDB.

B. Big Data Analysis and Mining

The content of data mining consists of six parts: correlation analysis, regression analysis, data classification, data clustering, data prediction and data diagnosis.

1) Machine Learning: The purpose of machine learning is to make a model by letting the machine observation data, and predict or explain the unobserved data. Machine learning is also the core of business insight and optimization analysis. Machine learning is now widely used in many fields. For example, when you are shopping on the Internet, you can use the recommendation system to push what you might want based on your historical purchase record. Also, Mailbox can automatically identify spam. Current machine learning systems are Spark [35], Petuum [36], TensorFlow [37] and Angel [38].

2) Statistical Analysis: Statistics is the science of collecting, organizing and managing data and drawing conclusions from it. Statistical analysis focuses on modeling, establishes reasonable models through probability, statistics and discretization, to mining data contents fully. Statistical analysis has two roles: one is to make inferential statistics, to model the data in some way, and to deduce the possible conclusions in the statistical data. The second is the statistical hypothesis test. After establishing or predicting the model, statistical verification is performed through the experimental data. If the result is unlikely to occur randomly, it is called statistical significance.

3) Data Visualization: After statistical analysis and mining of the data, there will be some understanding of the data, but not intuitive. The most intuitive way to understanding data is to visualize the data features so that you can understand the data more clearly and find the data connections that are not easily observed in the raw data. Common visualization techniques can be divided into two categorie based on whether the source data includes physical data. They are scientific visualization and information visualization.

C. Big Data Analysis and Processing Systems

The Big Data analysis and processing system can be divided into batch processing system, stream processing system, interactive processing system, graphic computing system and memory computing system according to the different types of analysis.

1) Batch Processing System - Google MapReduce:

2) Stream Processing System - Apache Storm:
3) Interactive Processing System -Apache Drill,SpagoBI:
4) Graphic Computing System -Pregel, Graph:
5) Computing System -Dremel,SolidDB:

IV. Pathologies of Big Data

Here provide an overview of a few of the issues that can arise when analyzing big data: the inability of many off-the-shelf packages to scale to large problems; the paramount importance of avoiding suboptimal access patterns as the bulk of processing moves down the storage hierarchy; and replication of data for storage and efficiency in distributed processing. Big data should be defined at any point in time as "data whose size forces us to look beyond the tried-and-true methods that are prevalent at that time." In the early 1980s, it was a dataset that was so large that a robotic "tape monkey" was required to swap thousands of tapes in and out. In the 1990s, perhaps, it was any data that transcended the bounds of Microsoft Excel and a desktop PC, requiring serious software on Unix workstations to analyze. Nowadays, it may mean data that is too large to be placed in a relational database and analyzed with the help of a desktop statistics/visualization package—data, perhaps, whose analysis requires massively parallel software running on tens, hundreds, or even thousands of servers. The big truth about big data in traditional databases: it's easier to get the data in than out. Most DBMSs are designed for efficient transaction processing: adding, updating, searching for, and retrieving small amounts of information in a large database. Data is typically acquired in a transactional fashion: imagine a user logging into a retail Web site , searching for products and making a purchase . A fair amount of data has been added effortlessly to a database that—if it's a large site that has been in operation for a while—probably already constitutes "big data." The pathologies of big data are primarily of analysis. This may be a slightly controversial assertion, but argue that transaction processing and data storage are largely solved problems. Short of LHC-scale science, few enterprises generate data at such a rate that acquiring and storing it pose major challenges today.In business applications, at least, data warehousing is ordinarily regarded as the solution to the database problem.A data warehouse has been classically defined as "a copy of transaction data specifically structured for query and analysis," and the general approach is commonly understood to be bulk extraction of the data from an operational database, followed by reconstitution in a different database in a form that is more suitable for analytical queries .To understand how to avoid the pathologies of big data, whether in the context of a data warehouse or in the physical or social sciences, we need to consider what really makes it "big."What makes most big data big is repeated observations over time and/or space.The fact that most large datasets have inherent temporal or spatial dimensions, or both, is crucial to understanding one important way that big data can cause performance problems, especially when databases are involved.In designing applications to handle ever-increasing amounts of data, developers would do well to remember that hardware specs are improving too, and keep in mind the so-called ZOI (zero-one-infinity) rule, which states that a program should "allow none of foo, one of foo, or any number of foo. That is, limits should not be arbitrary; ideally, one should be able to do as much with software as the hardware platform allows.Of course, hardware—chiefly memory and CPU limitations—is often a major factor in software limits on dataset size. Many applications are designed to read entire datasets into memory and work with them there. Any given computer has a series of absolute and practical limits: memory size, disk size, processor speed, and so on. When one of these limits is exhausted, we lean on the next one, but at a performance cost: an in-memory database is faster than an on-disk one, but a PC with 2-GB RAM cannot store a 100-GB dataset entirely in memory; a server with 128-GB RAM can, but the data may well grow to 200 GB before the next generation of servers with twice the memory slots comes out.

V. Implementation of the Big Data

Increasing amounts of data are streaming into contemporary organizations as a result of the rapidly growing quantity of data being generated not only by the organizations themselves but also in the organizations' business environments by both their stakeholders and other entities operating there. Thus, it is in this context that such expressions as "a data-centric world" have become more and more common . I The above mentioned processes are significant elements of the socio-economical changes taking place worldwide, where the extremely dynamic development of increasingly powerful and pervasive information technology has an important role to play. Advancements in this field have been a significant catalyst for the transformation of the contemporary economy and the emergence of an "interconnected economy".
This new type of economy, in the resource dimension, is a knowledge-based economy, where the most meaningful form of capital is intellectual capital . Under these conditions, from an organization's point of view, the ability to collect the right data and information and to transform it effectively into useful knowledge becomes an increasingly important issue. Recently, the field of information technology has begun to enter into a new era, as a result of the intensification of the progress being made there. It is an era where processing power and data storage have become virtually free, while networks and cloud based solutions provide users with global access and pervasive services. As a result of these processes, Big Data sets are being generated which have grown exponentially in size.In 2012, about 2.5 exabytes of data were created every day, with this amount doubling about every forty months .

A. Big Data and the most important sources of the growth of data

The term Big Data is not universally understood and applied, leading to various approaches to analyzing it. According to the Leadership Council for Information Advantage, this term is not precise "(. . .) it's a characterization of the never ending accumulation of all kinds of data, most of it

unstructured. It describes data sets that are growing exponentially and that are too large, too raw or too unstructured for analysis using relational database techniques" . On the other hand, NewVantage Partners describes Big Data as "a term used to describe data sets so large, so complex or that require such rapid processing (sometimes called the Volume/Variety/Velocity problem), that they become difficult or impossible to work with using standard database management or analytical tools" . It is important to underline that Big Data not only relates to the storage and consumption of original content but also to the data connected with this consumption . Generally, there have been some significant trends that have caused a considerable increase in data generation .

The first trend, the growth in traditional transactional data bases is chiefly connected with the fact that organizations are collecting data with greater granularity and frequency. This is due to various reasons such as the increasing level of competition, increasing turbulence in the business environment and the growing expectations of customers. All of these factors require organizations to react rapidly and with maximum flexibility to the changes taking place and then adjust to them. In order to be able to do this, they are forced to conduct more and more detailed analysis concerning marketplaces, competition and the behavior of consumers . The second trend, the increase of multimedia content, is connected with the rapid increase in the use of multimedia in the various industries of the contemporary economy, such as the health care sector where over 95 percentage of the clinical data generated is in video format.

B. Opportunities and Benefits Connected with Big Data Utilization

The development of the Big Data phenomenon and its associated tools and techniques, is not something which has been separated from the wider processes which have been taking place in organizations over recent years. In fact, it is becoming more and more common in organizations concerned with the field of analytics and has significantly expanded the possibilities available within the scope of business intelligence (BI) tools. Given their role in providing organizations with numerous possibilities and opportunities in the sphere of analytics, business intelligence systems are well suited for aggregating and analyzing structured data .

These mainly relate to situations where data sets become increasingly diverse, more granular, real-time and iterative. Such types of unstructured, high volume, fast-changing data, pose problems when trying to apply traditional approaches based on relational database models. There are many diverse benefits arising from the utilization of Big Data, depending on the sector of the economy, as has been confirmed by the results of research conducted by the McKinsey Global Institute. These results show the transformational potential of Big Data in such diverse domains as health care, public sector administration, retail, manufacturing and personal location data . An extremely important new element, in the context of decision-making, connected with the Big Data phenomenon, is the possibility for constant business experimentation to guide decisions and test new products, business models, and customer-oriented innovations.

According to the McKinsey Global Institute, five key ways in which Big Data creates value for organizations can be distinguished :

- Creating transparency by integrating data and making it more easily accessible to all relevant stakeholders
- Enabling experimentation to discover needs, expose variability, and improve performance
- Segmenting populations in order to customize actions,
- Replacing or supporting human decision making with automated algorithms,
- Innovating new business models, products and services.

Decision making based on data and business analytics also impacts on other performance measures such as asset utilization, equity return and market value .

C. The Most Significant Obstacles in Big Data

"Organizational silos" were the most significant barrier (55,7 percentage), which result from the fact that data connected with particular organizational functions (i.e. The second (50,6 percentage), although no less important, issue is the lack of appropriately skilled people (data scientists) prepared to analyze data. The third aspect (43,7 percentage) is the excessively long time it takes organizations to analyze huge data sets. As was mentioned earlier, organizations expect to be able to analyze and act on data in real time. The fourth barrier (41,7 percentage) is the difficulties concerned with the analysis of ever increasing amounts of unstructured data. Finally, the inability of senior management to view Big Data in a sufficiently strategic way (34,9 percentage) is the fifth key impediment .

VI. Big Data Challenges

A. Secure Computations in Distributed Programming Frameworks

Distributed programming frameworks utilize parallelism in computation and storage to process massive amounts of data.

B. Security Best Practices for Non-Relational Data Stores

Non-relational data stores popularized by NoSQL databases are still evolving with respect to security infrastructure. For instance, robust solutions to NoSQL injection are still not mature. Each NoSQL DBs were built to tackle different challenges posed by the analytics world and hence security was never part of the model at any point of its design stage. Developers using NoSQL databases usually embed security in the middleware. NoSQL databases do not provide any support for enforcing it explicitly in the database. However, clustering aspect of NoSQL databases poses additional challenges to the robustness of such security practices.

C. Secure Data Storage and Transactions Logs

Data and transaction logs are stored in multi-tiered storage media. Manually moving data between tiers gives the IT manager direct control over exactly what data is moved and when. However, as the size of data set has been, and continues to be, growing exponentially, scalability and availability have necessitated auto-tiering for big data storage management. Auto-tiering solutions do not keep track of where the data is stored, which poses new challenges to secure data storage.

D. End-Point Input Validation/Filtering

Many big data use cases in enterprise settings require data collection from many sources, such as end-point devicesA key challenge in the data collection process is input validation: how can we trust the data? How can we validate that a source of input data is not malicious and how can we filter malicious input from our collection? Input validation and filtering is a daunting challenge posed by untrusted input sources, especially with the bring your own device (BYOD) model.

E. Real-time Security/Compliance Monitoring

Real-time security monitoring has always been a challenge, given the number of alerts generated by (security) devices. These alerts (correlated or not) lead to many false positives, which are mostly ignored or simply "clicked away," as humans cannot cope with the shear amount. This problem might even increase with big data, given the volume and velocity of data streams.

The challenges posed by data volume is most visible. Scientists catch the computing limitation constantly due to the varying data volume. Web applications produce huge amount of data because they want to manage or they want to deal with lots of customers. These type of huge data can be seen in many areas like finance, business section, and in online transactions also. Regardless of these huge volume issues, there is no agreement on the quality of big data. These quality depends on various factors. Complexity of the data structure is the first important factor. A relational dataset is not referred to as a big data, because it can handled by today's DBMSs. But a graph dataset is referred to as big data. Because graph processing is very challenging to our technologies. Requirements of target applications should be considered as another factor. The quality of big data depends upon the requirements of target applications. Requirements includes response time, throughput etc. Some applications accept certain response time and some are very sensitive to the delays.

Velocity refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data. Big data velocity deals with the speed at which the data flows in from sources like networks, applications and mobile devices etc. The flow of data is massive and continuous. The challenge of velocity comes with the need to handle the speed with which new data is created or existing data is updated. In such applications where we want to handle huge amount of data, we require the systems to make sense of data immediately upon its creation. The technology of data streaming has been investigated for several years to handle high velocity. However, the capacity of the existing streaming systems is still limited, especially when dealing with the increasing volume of incoming data in today's sensor networks, telecommunication system, etc. In real-world applications, data often does not come from a single source. Big data implementations require handling data from various sources, in which data can be of different formats and models. This bring forth the challenge of data variety. The variety of data provides more information to solve problems or to provide better service. We can divide the data to capture the different types of data in a way that makes it possible to correlate their meanings. Typically, data can be classified into three general types—structured data, semistructured data and unstructured data. . Moreover, the challenges vary in different application scenarios.

Next challenge in big data variety, we know different applications are able to produce different types of big data, with many of them are unstructured and semi-structured. Heterogeneity is a natural property of the big data, considering their broad range of sources, we can divide the typical sources of big data as,

1) User Generated Contents (UGCs) : from applications with huge users. Examples of these big data are tweets, blogs, discussions, photos/videos posted and shared by users of many Web applications. The data of these applications are directly contributed by users, and therefore, they are typically unstructured for user convenience.

2) Transactional Data: that are generated by a large scale system due to massive operations/transactions processed by the system. Examples of big transactional data are Web logs, business transactions, feeds of moving objects, reports of sensor networks, reads of radio-frequency identifications. These data are typically structured with predefined designs. They are often accumulated in a streaming manner.

3) Scientific Data: that are collected from data-intensive experiments or applications. Examples are high-energy physics data, genome data, healthcare data. Types of scientific data are very applicationdependent, ranging from structured data (e.g., time series data) to semi-structured data (e.g., XML data) and unstructured data (e.g., images). In addition to the original data, provenance data (recording how data are generated and transformed) are very important for scientific data management.

4) Web Data: that are crawled and processed to support applications such as Web search and mining. As the World Wide Web contains billions of pages, it is quite easy to generate a huge unstructured Web pages. Behind the Web pages, there are also a huge amount of deep web data which are even more important than the surface contents.

5) Graph Data: that are formed by a very huge number of information nodes, and the links between the nodes. Examples are social networks Although structured, graph data are more expensive to process than relational data because pattern in graphs complicates the processing of graph data.

VII. Big Data Security

There are four different aspects of Big Data security: infrastructure security, data privacy, data management, and integrity and reactive security.

A. Infrastructure Security

It is mainly highlight the main technologies and frameworks found as regards securing the architecture of big data system and there are based on Hadoop technology, since it is that most frequently used.

1) Security for Hadoop: It is an opensource java based framework used for storing and processing big data. It is a distributed file system that enables concurrent processing and fault-tolerance. Hadoop stores large volumes of data and also it store and retrieve the data faster. So, the security problems related to Hadoop is a serious factors. In order to protect it, the solutions use different security mechanisms such as authenticity or cryptography.

2) Authentication: The value of the data obtained after executing a Big Data process be determined by its authenticity.The problem of authentication by creating an identity-based signcryption scheme for Big Data.

B. Data Privacy

It is probably the topic about which ordinary people are most concerned. A big data system usually contains an enormous amount of personal information that organizations use in order to obtain a benefit from that data.So we want to put some limits regarding the use of that information. Organizations should not have total freedom to use that information without our knowledge, although they also need to gain some benefit from the use of that data. Several techniques and mechanisms with which to protect the privacy of the data, and also allow companies to still make a profit from it.

1) Cryptography: Cryptography is one of the famous solution to securing data privacy in big data system. It is used to protect data for considerable amount of time. One example of the use of cryptography can be found in, bitmap encryption scheme that guarantees users' privacy.

2) Access Control: Access control is one of the basic traditional techniques used to achieve the security of a system. Its main objective is to restrict non-desirable users' access to the system. In the case of Big Data, the access control problem is related to the fact that there are only basic forms of access control. In order to solve this problem, there is a framework that supports the integration of access control features.

3) Confidentiality: Although privacy is traditionally treated as a part of confidentiality, they propose new techniques such as computing on masked data (CMD), which improves data confidentiality and integrity by allowing direct computations to be made on masked data.

1) Integrity: Big Data is supported is the capacity to receive streams of data from many different origins and with

C. Data Management

This section focuses on what to do once the data is contained in the Big Data environment. It includes the security of the data that stored in big data environment and also it includes how to share that data.

distinct formats: either structural data or non-structural data. This increases the importance of checking that the data's integrity is good so that it can be used properly. It protect data from un authorized alterations during its lifecycle.

2) Attack Detection: Big Data may be attacked by malicious users. Researchers develop a computational system that captures the provenance data and also there is a intrusion detection system especially intended for the specific characteristics of a Big Data environment.

VIII. Conclusion

This article describes the definition and characteristics of Big Data. In addition, according to the process of Big Data processing, it introduced the technical and processing tools.The rapidly growing amount of data which organizations have at their disposal and the opportunities connected with its practical utilization are increasingly changing the processes relating to making decisions at various organizational levels. Thus, Big Data offers huge potential to positively impact on the functioning of organizations generally and gives them a competitive advantage. But if initiatives aimed at the practical usage of Big Data sets are to be successful at giving an organization a competitive advantage and be of value, it is not enough to just collect and own the appropriate data sets.Security and privacy issues are magnified by velocity, volume, and variety of big data, such as large-scale cloud infrastructures, diversity of data sources and formats, streaming nature of data acquisition and high volume inter-cloud migration. The use of large scale cloud infrastructures, with a diversity of software platforms, spread across large networks of computers, also increases the attack surface of the entire system.
Traditional security mechanisms, which are tailored to securing small-scale static (as opposed to streaming) data, are inadequate. In the area of privacy protection, it introduced the different technologies of Big Data in releasing, storing and mining stages. Finally, this article predicts the trends of Big Data processing technologies.

References

[1] Yang Zhao-hong*A Literature Review on the Key Technologies of Processing Big Data*, The 3rd IEEE International Conference on Cloud computing and Big Data Analysis,2018
[2] Zhao Bin , A Literature Review on the Key Technologies of Processing Big Data,The 3rd IEEE International Conference on Cloud computing and Big Data Analysis,2018
[3] Wang Hui -yu ,A Literature Review on the Key Technologies of Processing Big Data,The 3rd IEEE International Conference on Cloud computing and Big Data Analysis,2018
[4] Han Zhi-he ,A Literature Review on the Key Technologies of Processing Big Data,The 3rd IEEE International Conference on Cloud computing and Big Data Analysis,2018
[5] Alvaro Cardenas Mora, Fujits,Top Ten Big Data Security And Privacy Challenges,Cloud Security alliance,CSA,2012
[6] Adam Fuchs, Sqrrl,Top Ten Big Data Security And Privacy Challenges,Cloud Security alliance,CSA,2012
[7] Jinchuan Chen ,Big Data challenge: a data management perspective,2013
[8] Xuan Zhou,Big Data challenge: a data management perspective,2013
[9] P.Joseph Charles,International Research Journal of Engineering and Technology (IRJET),Big Data Security an Overview,2018

Study on Cyber Security

Athira P V, Dhilsha K P and Sandra Thomas
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Salkala K S
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: salkala@vidyaacademy.ac.in)

***Abstract*—Cyber security is the collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment and organization and user's assets. This paper explore the definition of Cyber Security, how the term cyber security is often used interchangeably with the term information security ,the rules and guidelines to protect cyberworld and technologies that can used in cybersecurity.**

***Index Terms*—Information security, Risk, Cyber security, Threat, Blockchain, IoT, Artificial Intelligence**

I. Introduction

Cyber security has become a matter of global interest and importance. Cyber security is not necessarily only the protection of cyberspace itself, but also the protection of those that function in cyberspace and any of their assets that can be reached via cyberspace. Cybersecurity strives to ensure the attainment and maintenance of the security properties of the organization and user's assets against relevant security risks in the cyber environment. The general security objectives comprise the following:

- Availability
- Integrity, which may include authenticity and non-repudiation
- Confidentiality

With the advancement of computer and information technology, cybercrime is now becoming one of the most significant challenges facing law enforcement organizations

II. From Information Security to Cybersecurity

The term cyber security is often used interchangeably with the term information security. The definition of information security in depth and then argue that the boundaries of cyber security as a concept are wider than those of information security in terms of how it is formally defined. Cyber security goes beyond the boundaries of traditional information security to include not only the protection of information resources, but also that of other assets, including the person him/herself. Information security is the preservation of the confidentiality, integrity and availability of information. Cyber security as a protection of cyber space, the electronic information, the ICTs that support cyberspace, and the users of cyberspace in

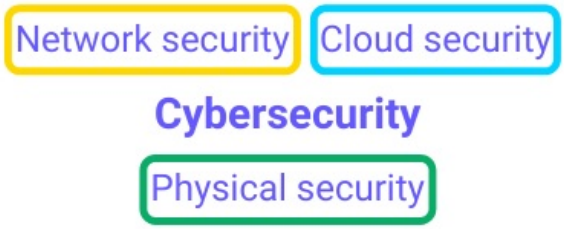


Fig. 1. Cybersecurity

their personal, societal and national capacity, including any of their interests, either tangible or intangible, that are vulnerable to attacks originating in cyberspace. It is clear that cyber security is far more extensive than any of the information and/or ICT security. The aim of information security is to ensure business continuity and minimise business damage by limiting the impact of security incidents Information security can be defined in a number of ways. In the case of information security, the assets to be secured is the information together with the underlying technologies. However, in the case of cyber security, the goal is clearly not to secure cyberspace but rather to secure those that function in cyberspace, whether individuals, organisations or nations.

III. Cybercrimes

The crime that involves and uses computer devices and Internet, is known as cybercrime. Cybercrime can be committed against an individual or a group; it can also be committed against government and private organizations. It may be intended to harm someone's reputation, physical harm, or even mental harm. Cybercrimes are increasing day by day. It always happens in the digital or virtual world. . In the physical world, crimes are unavoidable but can be minimized by all kinds of mechanisms, such as regulations, laws, legislation, police, and so forth. The advance of computers and networks has greatly contributed to the spread of cybercrimes. According to the main roles of computers or networks played in the crimes, cybercrimes are categorized into the following five classes.

1) The computer or network is used as a tool in a criminal activity.

2) The computer or network is the target of a criminal activity.
3) The computer or network is the place of a criminal activity.
4) Traditional crimes facilitated through the computers or the networks.
5) Other information crimes.

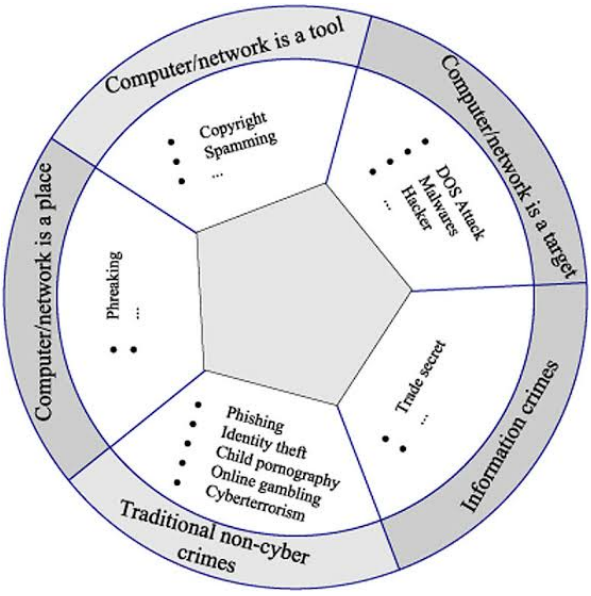


Fig. 2. Cybersecurity

A. As a tool

1) Copyright: Copyright is a legal concept, enacted by most governments, giving the creator of an original work exclusive rights to it with a limited time." Broadly speaking, it is "the right to copy", and the copyright holder achieves the right to be credited for the work, to decide who can use the work, who can perform the work, who can benefit from it, and other corresponding benefits.

2) Spamming: Spamming usually refers to the abuse of electronic messaging systems and the indiscriminate sending of unsolicited bulk messages. Spamming is widely recognized as e-mail spams; however, it has been similarly abused in other approaches such as instant messaging, the Usenet newsgroup, Web search engines, blogs, wikis, mobile phone messaging, Internet forums, and junk fax transmissions.

B. As a target

1) Denial of service: A DoS attack or distributed DoS (DDoS) attack is a crime that renders computers or network resources inaccessible to their intended users or customers. Although DoS attacks may be via different means, motives, and targets, they generally include the concerted, malevolent efforts of a person or persons to make an Internet site or service unable to perform normally or even at all.

2) Malwares: Malware refers to software designed to penetrate or destroy a computer system without the knowledge of the owner. The word malware combines the words malicious and software . As generally used by computer professionals, the expression refers to all kinds of software or program codes with hostile or intrusive purposes.

3) Hacker: In the context of security, a hacker is someone who tries to explore systems or obtain unauthorized access to others' computers through specific skills or knowledge. There are usually three kinds of hackers: black hat hackers, white hat hackers, and gray hat hackers. What people always refer to with the term "hacker" is a black hat hacker that is malicious or criminal. White hat hackers are ethical hackers, and those ambiguous in ethics are called gray hat hackers.

C. As a place

1) Phreaking: Phreaking is a slang term used to describe criminal activities related to phones. The word "phreak" is a combination of the words "phone" and "freak", and those individuals involved in "phreaks" are known as "phreakers" or "phone preaks". Such criminal activities are always related to those people working with or studying telecommunication systems, especially those familiar with public telephone networks and related equipments and systems.

D. Traditional non-cybercrimes facilitated by computer and network technology

1) Phishing: In computing, phishing refers to attempts to criminally and fraudulently gain sensitive information, such as usernames, passwords, and credit card details, by means of some public entities that run on electronic systems, such as online banks, PayPal, and eBay.

2) Identity Theft: Identity theft is a term used to describe fraud in which the criminal pretends to be someone else to steal money or get other benefits. It is also a crime for criminals to pretend to be someone else even if they do not steal an identity.

3) Child pornography: Child pornography is the term used to describe the sexual abuse of children by means of pornographic material. With the help of the Internet, it is quite easy to spread images and video. Child pornography is illegal all over the world. Related production of such material is also prohibited. The main reason that such criminal activities continue is the profit that can be generated from the sale of such images.

4) Online gambling: Online gambling generally refers to gambling over the Internet.

5) CyberStalking: As defined in, cyber stalking is stalking someone through the Internet or through other electronic means. It refers to an individual or group of individuals harassing another individual, group of individuals, or organization through the Internet or other communication technologies.

6) Cyber terrorism: With the intense development of information technology, the Internet has become more pervasive in all areas of human life. Under the threat of crimes committed in cyber space, individuals, social organizations, or even the

whole Nation may be attacked. Cyber space attackers do not face the inherent threat of injury or death that a physical attack would bring.

E. Information crimes

1) Trade secret: A trade secret is considered by a business to be some advantage over competitors or customers. The advantage can be in many different forms, such as formulas, practices, processes, designs, instruments, patterns, or compilations of information. In some jurisdictions, terms such as "confidential information" or "classified information" are also used to describe such secrets.

IV. Cyber Security Attacks on Smart Cities and Associated Mobile Technologies

There is no specific and universal definition for the concept smart city as countries and their governments determine how far they go for smartening according to their willingness to change, resource limitations and financial status. There are five main components that are essentially required to be in a smart city: modern information and communication technologies, buildings, utilities and infrastructure, transportation and traffic management and the city itself. smart cities aim to secure economic competitiveness in urban spaces and let urban citizens experience classier lifestyles. Clearly, smart cities are complex and interdependent so it face some problem mainly security and privacy. Insuring security in smart cities indicates protecting data, information and network from any attack and malicious activity. Deployment of hardware and software usually released by vendor without specific testing for cyber security. By using such things leads hacking, which the system fill with fake data and caused system failure. Smart cities provide huge benefits to user, but at the same user concern about data privacy. Users must pay huge attention to what they share and must know that once they share any piece of personal data, it will not disappear. It is important to alert planners and analysts for the necessity of thinking about protection against security vulnerabilities during the design of smart city. Security in smart city:

A. Infrastructure security

Intentional and accidental threats on smart city infrastructure-related security cause different serious consequences according to city's maturity and smartness. Thus, we present main encountered infrastructure security-related threats and challenges. Urban infrastructure such as electricity supply, water distribution, streets, buildings and others face several security threats in their specific cyber-physical components and systems such as: Cameras, communication network, building management system, transport management systems.

B. Data / information privacy in smart cities

Smart Cities deal with huge quantities of real time data and related technologies that is data-driven technologies that act on, generate, process, run, and produce data. Smart Cities have many resources producing different types of data. Among these resources are systems that continuously produce fine-scaled and exclusive data. These systems are widespread in smart cities and the data they produce are called big data. All these urban data are used to run smart city technologies, so it is important to keep these vast amounts of data and information secure. Moreover, it is necessary to maintain the privacy of locked data and personal data and to make sure that they are not intentionally or accidentally reached or accessed. Privacy is ensured by protecting five privacy related issues: protecting identities that indicate protecting personnel and their confidential data; protecting people areas that indicate to protect each one's space and properties; protecting locations which indicate preventing spatial tracking; communication protection which indicate not to eavesdrop any kind of conversations; and finally, transactions protection that protect every single purchase, exchange and query.

V. Ten Rules for Cybersecurity

The spectrum of cyber conflict ranges from breaches of internal policy or regulations to breaches of legal obligations to crime to national-security threats to out-right cyber warfare. The levels and sources of law relating to cyber security range from the soft to organisational to national to international agreements and customary law, which inform the four key legal areas the law of network and information security, dealing with, for example, data protection , e-commerce, electronic communications and access to information; criminal law (offences, investigation, cooperation); national -security law and possible restrictions to human rights and liberties resulting from national-security concerns; and the law of armed conflict. Contemporary cyber threats can only be confronted by combining the regulation, remedies and legal practice these four key areas of law. Ten rules focused on issues and working solutions arising from discussions among experts or in the course of cyber-incident handling can be identified.

A. The Territoriality Rule

Cyber infrastructure is subject to the jurisdiction of the flag state and is subject to the sovereign prerogatives of that state. Every government can exercise effective control over the IT infrastructure located on its territory, for example by ensuring the availability and quality of logs, maintaining an overview of the providers of electronic communications, developing an understanding of threats and capabilities existing within its jurisdiction to cope with and manage incidents, and balancing the development of the information society with the interests of national security.

B. The Responsibility Rule

Information leading to the identification of the source of an attack or about the perpetrators, methods and tools involved, and even active law enforcement measures such as confiscation, arrests and prosecution, should reasonably be expected from the countries whose infrastructure has been involved.

C. The Cooperation Rule

The interconnectedness of global information infrastructure makes it impossible for any nation to defend itself against cyberattack without cooperating with states whose infrastructure can be used to route such an attack. More effective cooperation is needed between public and private institutions as well as between national governments and international organisations. Cross-disciplinary cooperation between legal, policy, military and technical experts is also necessary.

D. The Self-Defence Rule

The concept of self-defence is part of both criminal and international law. In principle, everyone has the right to self-defence, subject to the proportionality and necessity of such actions.

E. The Data Protection Rule

The need for network monitoring and information exchange has to be carefully assessed against individuals' right to privacy. There is currently a considerable divide between the legal and technical approaches to data and their security. While the monitoring of network data seems to be well-established and routine in technical communities, it raises significant concerns among legal experts.

F. The Duty of Care Rule

The concept of duty of care is well established in many areas of law: an individual is under obligation to guarantee the protection of personal data he processes, and due-diligence duties arise from the legal framework of data protection, information-society services, consumer protection and so on.

G. The Early Warning Rule

The attack itself involved a single, easily fixed ISP vulnerability, but the response had broader ramifications: having learned of the upcoming attacks, the ISP issued an early warning to its customers and informed them about the incident. If implemented widely, this approach could considerably improve cyber security.

H. The Access to Information Rule

There is a strong trend in Europe towards transparency of governmental acts and records, giving the public the right to be informed about threats and decisions related to their life and well-being. A holder of information is required to disclose existing information to danger to the life, health and property of persons. The presumption is that public-sector information should be publicly accessible unless there are compelling reasons otherwise. While access to information can allow the public to learn of threats and attacks and can raise awareness about cyber security, it may also result in unwanted publicity.

I. The Criminality Rule

The criminality rule is a reminder rather than something qualitatively new. It is well established in criminal law that cyber attacks can only be investigated and prosecuted if those acts qualify as criminal offenses.

J. The Mandate Rule

The mandate rule is relevant for defining and coordinating international efforts in global cyber security. Its particular, practical importance lies in the realm of developing new or revising existing cyber-security agendas.

VI. Blockchain In Cyber Security

As a cryptographic-based distributed ledger, blockchain technology enables trusted transactions among untrusted participants in the network. Since the introduction of the first Bitcoin blockchain in 2008, various blockchain systems, such as Ethereum and Hyperledger Fabric have emerged with public and private accessibility outside of existing fiat currency and electronic voucher systems. Recently, blockchain technology has also been the subject of increased scientific research and development and has raised significant interest among researchers, developers and industry practitioners due to its unique trust and security characteristics. There is no doubt that the popularity of blockchain has increased worldwide. unique properties of the blockchain technology make its application an attractive idea for many areas of business, such as banking logistics the pharmaceutical industry smart contracts and, most importantly in the context of this paper, cyber security. Our overarching goal is to provide a community-driven initiation for a better study of blockchain and cyber security that explores the interplay between the two highly discussed fields. Toward this goal, we will critically examine existing works and studies on blockchain cyber security, and use these insights to develop new directions. The purpose of this research is to analyse existing studies, analyse their findings and summarize the efforts of research into blockchain applications for cyber security. To assist in focusing the work, we developed five research questions.

A. Research Questions

1) What are the latest security focused blockchain applications? : Use cases for blockchain have diversified away from solely cryptocurrency. A review of the latest practical applications will help with understanding the full impact of blockchain technology on cyber security.

2) How is blockchain used to improve cyber security?: Blockchain features can be deployed to solve problems relating to the security of devices, networks and their users. This will provide an understanding of the methods used to implement blockchain in digital infrastructure for the purpose of security.

3) What methods are available for blockchain solutions to manage security without requiring a cryptocurrency token? : Cryptocurrency blockchains are commonly maintained through a proof-of work mechanism whereby miners can show to the rest of the network that they have invested significant resources in order to assist in the validation of transactions. This question will look at research that addresses how a blockchain can be maintained without the requirement to incentivize miners for transaction validation.

B. Future Research Directions of Blockchain in Cybersecurity

Based on the results of this survey and our observations, we present the following research directions of blockchain in the field of cyber security that worth further investigation:

1) Blockchain in IoT Security: Security in IoT networks has been claimed as a pressing need of the industry that gets utmost priority for improvement and enforcement. Despite current research and the fact that almost any article on blockchain cyber security in the literature points out that the security of IoT systems could be revitalized if it is supported with blockchain technology. Yet, little is known and discussed about factors related to decisions and feasibility to adopt this technology, and how and where it can systemically be put into use to remedy current IoT security's risks/threats in a clear context, allowing to imagine and then create future vectors in this specific domain. Thus, it is important for future research to develop some kind of quantifiable guidelines and tools that can help realize this gap in the literature. Furthermore, proposing lightweight blockchain-based solutions for resource constrained IoT devices (running on the edge of network) could be another avenue of further research.

2) Blockchain for AI Data Security: In modern computing ecosystem, data is captured from various sources and is transmitted from devices (e.g., IoT) through the networks. Artificial intelligence (AI) and its derivatives have been used as powerful tools to analyze and process the captured data to achieve effective reasoning in addressing security issues. Although AI is powerful and can be engaged with distributed computing, deceptive analysis would be generated when corrupted or dishonest data is intentionally or unintentionally integrated by a malicious thirdparity based on adversarial inputs. Blockchain as a popular ledger technology has the potential to be leveraged in different aspects of cyber space. The blockchain attempts to reduce transaction risks and financial fraud, due to characteristics such as decentralization, verifiability and immutability for ensuring authenticity, reliability and integrity of data. When the trust and reliability of the data can be ensured, more secure and trustworthy outcomes can be produced by the AI. A future research direction could be the exploration of blockchain in security of AI data in B2B and M2M environments.

3) The Sidechain Security: The sidechain technology has most recently emerged as a separate chain attached to the main chain, in parallel with transactions, to alleviate the challenges related to main blockchains.

VII. The future of Artificial Intelligence in Cybersecurity

Cybersecurity characterized as a collection of processes that help protect electronic data, human activity, and systems. Artificial Intelligence or AI is the development of complex computer systems with the aid of human mentality and able to perform its function like a normal human being, for example, it can recognize the voice and processes it into different languages as a human being. AI is the comprehensive scientific system with varying branches in math, computer science, and philosophy that purpose to develop another intelligent system that shows intelligence properties usually, the word artificial intelligence mostly utilized to describe the machinery that emulates "perception" functions, which humans relate to their mind, i.e. problem- solving and swotting. Machine learning is an essential factor in modern research and business. Algorithms are used and network-neutral models to oblige the systems of computers to improve the successfully Development. Machine learning algorithms automatically build a mathematical model by using sample data which is called training data to make decisions without being specifically arranged. Machine learning is based on a model that is the interaction of the brain cell. Types of learning:

1) Supervised Learning
2) Unsupervised Learning
3) Reinforcement Learning

Block Chain Technology has investigated the existing and future business applications of particular accounting and cyber security [25]. For existing cyber and accounting problems, authors apply to block chain uses. Researchers [25] are reviewing literature that includes topics such as large-scale data in the accounts, use of the financial security and cyber security chain and use of ledger technology in financial accounts, and financial malfeasance tracking. s. Researchers show that the block chain has different auditing consequences that will change the profession drastically. Researchers also believe it necessary to effectively implement the block chain in a number of fields, such as auditing and accounting, of cyber security and accounting. Internet of Things (IOT) is a network of different internet-connected devices capable of collecting and exchanging data, and these IOT devices produce a great deal of information to be gathered and stored for use in artificial intelligence (AI) to handle enormous data flows and storage through the IOT network. The paradigm shift towards the Internet of Things (IoT) has created a huge capacity for different future IoT scenes such as smart home, intelligent transport, smart health and smart energy, and the advance of the edge measurement concept. There are also a host of emerging cyber-security threats. Artificial intelligence was just a sort of computerized human intelligence edition. The way artificial intelligence functions was like learning, as people do, iteratively over and over. The threat of landscape evolves unquestionably in this century. The cyber attackers are based solely on financial incentives. But the department has found a new way to deter attacks before they happen because it can't depend on the old conventional way anymore. This paper highlights the need to develop cyber safety skills and how the use of artificial neural networks and machine learning algorithms can mean improved skills. The overview and definition of social engineering, the role it plays in networking and cyber identity robbery, and the reasons and impact on cybercrimes are also included. The authors finally recommend preventive action and potential solutions to threats and vulnerabilities in social engineering. AI approaches will simplify and render the cyber security computing process smart over traditional security systems. The production of IT will make a machine

behave and think like people. AI is an exceptional aspect of IT which requires the development of a computer that reacts and functions as a human mind. The core aspects of artificial intelligence include the comparison of the human senses. The system is able to recognize touch and speech as features placed within the system for operating the possible activities of a normal life situation without human assistance. However, artificial intelligence is the intelligence agent research, which takes the state of the world and successfully achieves its target. The majority of computer structures in the world are constructed to serve the purposes according to the essence of the situation with the special application of the human elements. Artificial intelligence is usually a human being who uses troubleshooting methods and learning to understand the high levels of activities in the operation of elements inspired by man, decision making and the emotional cycle. Artificial intelligence is machine-based intelligence as opposed to human intelligence. The research paper aims to assess the emerging issues relating to artificial cyber security intelligence in the United States. The research paper proposes the groundbreaking Artificial Intelligence Cyber Security approach in the United States. Artificial intelligence methods have quickly advanced in recent years and their uses in many fields can be seen from face recognition to picture processing. AI technology will improve cyber security tools and allow opposes to evolve attack methods in the cyber security market. But malicious players are aware of the new prospects and are likely to try to hurt them. cybercrimes are becoming increasingly complex, cyber security approaches are needed to be more robust and intelligent. This will allow defence mechanisms to make real-time decisions that can react effectively to sophisticated attacks. To help this, researchers and practitioners need to know existing cyber-security methods. The use of artificial intelligence in particular in the fight against cybercrimes. However, artificial intelligent approaches to combat cybercrimes are not summarised. The papers were reviewed using quantitative and qualitative approaches using a systematic mapping in the fields of IoT, Blockchain, Cybercrimes, Business, IDS, Software defined networks and cyber forensics. It was found that artificial smart methods contributed remarkably to cybercrimes by significantly improving intrusion detection systems. It was also found that computer complexity, model training times and false alarms have been reduced. However, the domain is significantly skewed. Most research centred on intrusion detection and prevention systems and support vector machines were the most dominant technique used.

VIII. Conclusion

This paper summarizes cybersecurity and its different aspects. The paper argued that cyber security, despite often being used as an analogous term for information security, differs from information security and it describes cybercrimes and the rules to ensure cybersecurity. Mainly cyber security is affected by the emergent integration of technologies and the resulted intensive communication, high complexity and high interdependency, which leads to unbounded attack surface and cryptography-related issues. This paper also has carved out relevant recent research available on how blockchain solutions can contribute to cyber security problems and an overview of how artificial intelligence can be used in the sense of cyber security in both crime and defence.

References

[1] Feng Tao , Muhammad Shoaib Akhtar and Zhang Jiayuan," The future of Artificial Intelligence in Cybersecurity: A Comprehensive Survey"
[2] Anwaar AlDairi and Lo'ai Tawalbeh "Cyber Security Attacks on Smart Cities and Associated Mobile Technologies "
[3] Ten Rules for Cyber Security, Survival: Global Politics and Strategy, by Eneken Tikk
[4] Rossouw von Solms, Johan van Niekerk "From information security to cyber security"
[5] A Systematic Literature Review of Blockchain Cyber Security - Paul J Taylor
[6] A survey of cyber crimes - Y. Zhang
[7]) Petrolo R, Loscrì V, Mitton N. Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms. Transactions on Emerging Telecommunications Technologies. 2015 Feb 1.
[8] Z. Siddiqui, M. S. Husain, and S. Yadav, "Application of Artificial Intelligence in Fighting Against Cyber Crimes: a Review," Int. J. Adv. Res. Comput. Sci., vol. 9, no. 2, pp. 118–122, 2018.
[9] Dhillon G. Principles of information systems security. John Wiley and Sons; 2007
[10] Vanolo A. Smartmentality: The smart city as disciplinary strategy. Urban Studies

Survey of Machine Learning Applications

Devika P H, Shejila K, and Meghana K B
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Aparna S Balan
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: aparna@vidyaacademy.ac.in)

***Abstract*—Machine Learning is an interesting research area in computer science which changes the thoughts of human to machines. It is mainly used for data analysis that plays a vital role in many application areas where more data is generated and need to turn that data into useful information. Focusing on different application domains of machine learning algorithms, machine Learning (ML) has been enjoying an unprecedented surge in applications that solve problems and enable automation in diverse domains. Primarily, this is due to the explosion in the availability of data, significant improvements in ML techniques, and advancement in computing capabilities. There are various surveys on ML for specific areas. This paper provides a survey about different applications of machine learning in various areas. In this way, readers will benefit from a comprehensive discussion on the different applications of ML and algorithms applied to fundamental problems in traffic prediction , health care, spam detections, web attacks detection, malware detection, app recommendation , malicious application detection etc.**

***Index Terms*—machine learning,health sector, predictions, detections, recommendations,**

I. Introduction

Machine Learning (ML) is the computerized approach to analysing computational work that is based on both a set of theories and a set of technologies. Although machine learning is a field within computer science, it differs from traditional computational approaches. In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve. Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs. Machine learning applications are types of technologies that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so.The main objective of machine learning applications is to improve business decisions, increase productivity, detect disease, forecast weather, and do many more things. MHealth is an important mechanism for improving the self management of the asthma patients[1]. Mobile health (mHealth) offers a promising platform for combining passive and active approaches to deliver an engaging and effective self-management system for asthma.

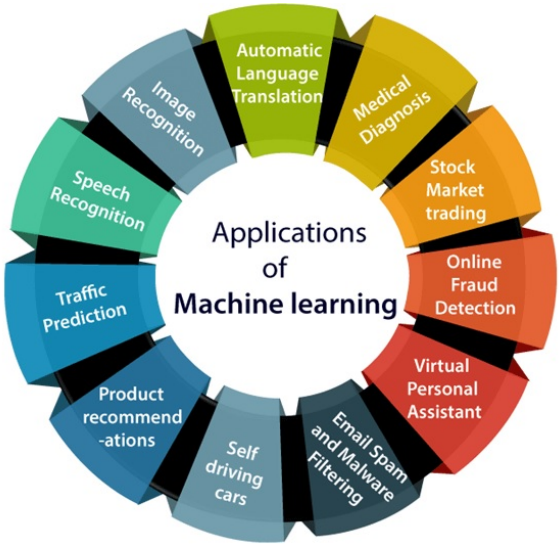


Fig. 1. Applications of Machine Learning

MyAsthma is a National Health Service (NHS) approved mobile app developed by mHealth[2]. Asthma MD has similar features, logging user asthma activity, peak flow ,medications and triggers, has paperless action plan storage, and can provide custom notifications[3]. It focus on helping patients and doctors complete their treatment tasks by using predictable test results based on the International Classification of Diseases (ICD-10) and helping hospitals save time and reduce effort dedicated to medical testing[4]. SEMLHI software engineering method also improving the efficiency of health system. It provides interaction between software engineering and machine learning within the context of health systems[5].Malicious application detection methods were used in machine learning are Spydroid, Dongfang, Mayank jayswal, etc. Another important application is predicting accurate and timely traffic flow Information. This prior information which is very near approximate about all daily life situations which can affect traffic then, a driver or rider can make an informed decision.ML is also helpful for recommending apps according to user's needs. So It helps to identify the malware apps and also ensure that if it is

safe or not. By developing a mobile application recommendation system with ratings, size and permission as parameters an it will recommend suitable apps for the users. Most of the web applications that are permanently being exposed to attacks that exploit their vulnerabilities. By the use of machine learning techniques to compute the performance of Web Application Firewalls systems (WAFs) that are used to detect and prevent attacks. Android platform due to open source characteristic and Google backing has the largest global market share. Being the world's most popular operating system, it has drawn the attention of cyber criminals operating particularly through wide distribution of malicious applications. Electronic mail has eased communication methods for many organisations as well as individuals. This method is exploited for fraudulent gain by spammers through sending unsolicited emails. Then it aims to present a method for detection of spam emails with machine learning algorithms that are optimized with bio-inspired methods. The main contribution of the work is reduction of feature dimension to less than half of original feature-set using Genetic Algorithm such that it can be fed as input to machine learning classifiers for training with reduced complexity while maintaining their accuracy in malware classification. In contrast to exhaustive method of feature selection which requires testing for 2N different combinations, where N is the number of features, Genetic Algorithm, a heuristic searching approach based on fitness function has been used for feature selection. The optimized feature set obtained using Genetic algorithm is used to train two machine learning algorithms: Support Vector Machine and Neural Network. It is observed that a decent classification accuracy of more than 94% is maintained while working on a much lower feature dimension, thereby, reducing the training time complexity of classifiers.ML applications included in some of the fields such as health sector, software engineering sector, detections ,predictions.

II. mHEALTH

Asthma is a condition in which your airways narrow and swell and may produce extra mucus. This can make breathing difficult and trigger coughing, a whistling sound (wheezing) when you breathe out and shortness of breath. The dualfocus in the management of patients with asthma is on symptom control and the prevention of asthma attacks. There is no cure for asthma. However, existing management strategies, such as the use of preventer inhalers, can be used to control the condition. Supported self-management (including an action plan) significantly reduces the risk of an asthma attack. Mobile health (mHealth) offers a promising platform for combining passive and active approaches to deliver an engaging and effective self-management system for asthma. Here this method develop an effective and engaging mHealth system that facilitates self-monitoring and uses personalized algorithms to provide timely and appropriate feedback (early warning) to patients and alsoemployed a publicly available mHealth dataset to apply and benchmark machine learning techniques to develop early warning algorithms. For analyzing the data, The key steps are data pre-processing and labelling, feature extraction, feature selection, classification, and model evaluation.

1) The AMHS was conducted via Asthma Health App, an Apple app designed using Research Kit for the purpose of the study, and contains data collected.
2) Data Pre-processing and Labelling: analyses the subset of patients with at least one weekly survey entry and at least three daily survey entries with peak flow readings. Information about a patient's condition (stable or unstable) and symptoms were derived from the weekly and daily survey respectively.
3) Feature Extraction: feature extraction method allowed a direct comparison between periods of time despite the disparity in data availability and sampling frequency.
4) Feature Selection: The least absolute shrinkage and selection operator (LASSO) method was employed to rank and select a handful of input features. The regularisation performed in LASSO not only helps us avoid over- fitting but also helps in ranking input features based on their predictive power.
5) Classification: this corresponds to developing an algorithm that could classify whether a patient is in a stable or in an unstable condition. We used a number of well-known machine learning classifiers, both linear and non-linear, and both probabilistic and deterministic models. The models used in this study consisted of decision trees, logistic regression, naıve Bayes, and support vector machine (SVM).
6) Performance Metrics : The skewed nature of the data made the geometric mean accuracy (GMA) a more suitable metric than accuracy for identifying the optimal threshold to maximise the sensitivity and the specificity. A set of GMA can be evaluated from the points on the ROC curve, then the maximum of this set represents the best threshold for the given data.

III. SEMLHI

It provides interaction between software engineering and machine learning within the context of health systems. It provides novel framework for health informatics: the framework and methodology of software engineering for machine learning in health informatics (SEMLHI). The SEMLHI framework includes four modules as software, machine learning, machine learning algorithms, and health informatics data that organize the task sin the framework using a SEMLHI methodology, thereby enabling researchers and developers to analyze health informatics software from an engineering perspective and providing developers with a new road map for designing health applications with system functions and software implementations.It focus on helping patients and doctors complete their treatment tasks by using predictable test results based on the International Classification of Diseases (ICD-10) and helping hospitals save time and reduce effort dedicated to medical testing. Using the SEMLHI framework, realistic patient data were analyzed carefully and rigorously based on important parameters such as age, start time, end time, patient treatment,

and detailed treatment content for each task. It consist of the following steps:

- Dataset And Preprocessing: The SEMLHI models and methodology were developed by including new software systems connected to real datasets and presented knowledge from the data using ML algorithms to improve the efficiency of the required system.
- Available Features Of Patients: The patient dataset included 457914 cases and nine tables. Each table had different features, and many techniques could be implemented, such as semantic coordination for intelligent databases, feature selection problems using genetic algorithms, and new gene-weight mechanisms. Some features were connected with other tables to build datasets describing the main attribute. Some of the software frameworks and their functionalities;

	SEMLHI	Vee	Agile
Flexibility	very high	rigid	very high
Emphasis	risk	specification	customer
Logic	depth first	breadth first	depth first
Assumption	independent iteration	stable info	independent iteration
Scope	medium and large	large and complex	small
Iteration	very quick	slow	very quick
Delivery	one-shot delivery	one-shot delivery	incremental delivery

Fig. 2. Table 1:Comparison of the three system engineering methods, Vee, Agile and SEMLHI, proposed in this article

SEMLHI METHODOLOGY: The SEMLHI methodology is used in software development in the health area. This framework includes ten stages, starting from defining the problem until reaching the stage of development and ending with the results, as explained in the next section. To develop the HI system, developers follow many sequence steps, such as design, implement, maintain defined workflows, structure information, provide security and privacy, test the performance, and then release the software applications. Records in most datasets in HI are weakly structured and non-standardized. To apply ML to the HI system, a set of patterns must be used by the algorithm to predict and visualize the ML algorithm and generate knowledge. The main patterns that were used in our framework were the geographic location, patient records, departments and hospitals, surgical history, obstetric history, family history, habits, immunization, assessment and plan, and test results. SEMLHI Framework was composed of four components or modules (software, machine learning model, machine learning algorithm, and health informatics data).

1) Health Informatics Data: To use a dataset on health informatics data (HID) algorithms, a transformation into numerical features was required. Other data contain missing, duplicate or null values, such as negative ages and extremely large integers, which could negatively affect the performance of our ML algorithm.
2) Ml Algorithms: Machine learning algorithms (MLAs) are used to compute the parameters that might define a model, optimize its network topology and improve the system convergence without losing information.

Fig. 3. SEMLHI Framework

3) Machine Algorithm Model:The machine algorithm model (MAM) component includes five submodules: read the data, prepare the data, train the model, test and evaluate the model, and predict new data.
4) Software: The software module includes a subclass that includes reuse, performance, testing, privacy, and security.

IV. Malicious Applications Detection in Android

Malicious apps set out to harm a device or the data on the device. They often steal user data, commit financial fraud, negatively impact device performance, and more.Various kinds of attack can occur because the amount of detection will be limited. To overcome this, a technology is introduced as SIGPID (Significant permission identification). The goal of SIGPID is to improve the apps efficiency .This technique will uses random forest algorithm, decision trees and support vector machine.

Spydroid has been proposed for real time usage and will use many detectors which will be obtained from third parties and will be much more efficient and enable real time monitoring[6]. It consists of Spydroid monitoring module and android open source project. To obtain the information, the OS needs to be changed and special tools can be used if required. Spydroid consists of some modules of operating system and application layer and sub detectors are proposed.1) A Sub detector is a normal android application that will monitor the runtime information using a monitoring module Zinal D[7].2) Patel used madam detection algorithm. It have many drawbacks such as some techniques performed real time analysis and

monitored the user devices as the third-party apps were not provided low-level information by the android.

Some methodology for mobile analysis to detect malicious applications that deceive users by not showing some of their functions . They are:

1) Dongfang li uses deep neural network. It is a signature based method, it is also having drawback, because it is strongly depended on the finger prints database.[8]
2) Mayank Jaiswal used system call analysis to identify whether the applications were malicious or legit. Here the call frequency is also used to classify the game as malicious or not .Every apps has their own system call frequency. Sometimes the system get harmed due to (apps requested a very high level permission from users) by taking undue advantage of the permission granted[9].
3) The other method is participatory sensing network sensors .Here calibrations has been studied using sensing network which have uncontrollable participants. Sensors are then calibrated using a process named as monitoring pollution sources which doesn't utilize celebrative process .The number of sources is limited and are spread over a large area sensing range is negligible to the distance between pollutions also.

Malicious Application Detections mechanisms in machine Learning consist of: System architecture, Data collection, Preprocessing, Feature extraction , Efficiency calculation.

In system architecture it representing the dataset divided into training and test dataset. The training dataset is to be the input to the algorithm. After finding the predicted output compare it with the actual outcome for calculating the efficiency. Data collection mainly presenting the attributes with values 0 and 1. 0 for it is not asking for permission and vice versa and also a column that indicate whether the app is malicious or not. Pre-processing will be the removal of whole row or column according to the suitable adjustability. So, the data is cleaned and noisy data is removed from dataset. Feature extraction is mainly stating that datasets have more than one attributes, some of the attributes are extracted from the dataset and it is used to input to the algorithm for better efficiency. After the feature extraction is complete, the data is split into training and testing data. In the efficiency calculation after training the algorithm the test data is input is given as the input to algorithm and compare the predicted with actual results to find the efficiency of both algorithm individually. The results will represents in 3 graphs. We can observe that as the size of Training dataset increases the corresponding errors decrease.

V. TRAFFIC PREDICTION

Various Business sectors and government agencies and individual travellers require precise and appropriately traffic flow information. Traffic Environment involves everything that can affect the traffic flowing on the road, whether it's traffic signals, accidents, rallies, even repairing of roads that can cause a jam. By developing a tool for predicting accurate and timely traffic flow Information then the prior information which is very near approximate about all the above and many more daily life situations which can affect traffic then, a driver or rider can make an informed decision. The development and deployment of Intelligent Transportation System (ITSs) provide better accuracy for Traffic flow prediction. It is deal with as a crucial element for the success of advanced traffic management systems, advanced public transportation systems, and traveller information systems[11]. The dependency of traffic flow is dependent on real-time traffic and historical data collected from various sensor sources, including inductive loops, radars, cameras, mobile Global Positioning System, crowd sourcing, social media. Traffic data is exploding due to the vast use of traditional sensors and new technologies, and we have entered the era of a large volume of data transportation. Transportation control and management are now becoming more data-driven. However, there are already lots of traffic flow prediction systems and models; most of them use shallow traffic models and are still somewhat failing due to the enormous dataset dimension.[12].

Deep learning is a part of machine learning algorithms,. In DL, use concepts of a neural network, by using this feature, it is beneficial to find network dynamics (such as spectrum availability, congestion points, hotspots, traffic bottleneck[13]. The travel time is the essential aspect in ITS and the exact travel time forecasting also is very challenging. Vector Machine (SVM) is one of the most effective classifiers among those which are sort of linear. It is advantageous to prevent overfitting of data. SVM is great for relatively small data sets with fewer outliers[14].

A. Implementation

Different machine algorithms applied and tested for achieving higher efficiency and accurate results. To identify classification and regression Decision Tree Algorithm (DT) is used. The goal of this method is to predict the value of the target variables. Decision tree learning represents a function that takes as input a vector of attributes value and return a "Decision " a single output value. It falls under the category of supervised learning algorithm. It can be used to solve both regression and classification problem[15]. Outliers detection is another critical step for an accurate result, Support vector machines (SVMs) is used. The SVM is beneficial for high dimensional spaces, and it also helps in the condition where a number of samples are less than the number of dimensions. The random forest algorithm is based on forecasting models, and it is mostly used to classify the data DT identify its results by performing a set of tests on the training dataset[16]..

The proposed algorithm for identifying the congested situations, they collect data at every 5 min with features- Location, Direction, speed and Start -End junction. Group these data according to these 5 min interval. Then calculate the distance between the vehicle if it is less than specific threshold it is consider as neighbour hood otherwise not.

Steps Involved in implementation:-

1) Created the application which can provide us the GPS coordinates.
2) Perform the proposed algorithm

3) Evaluate the matrix for the dataset
4) Divide the dataset into training and testing.
5) Analyse different machine learning algorithms.
6) Predict the 45 min interval parameters through machine learning algorithm
7) Conclude about the traffic congestion.

On implementing this algorithm and it can obtain the model which gives higher accuracy of the machine learning than the existing ones. Deep learning and genetic algorithm is an important problem in data analysis, it has not been dealt with extensively by the ML community.

VI. Mobile App Recommandation System

Generally applications are developed to satisfy the different needs of users. Most of the applications are not secured. Some of them are under the category of malware which will make the system down or misusing personal details etc.So it is necessary to identify this malware apps and also ensure that if it is safe or not. By developing an Android App Recommendation System for recommending safe apps from the play store. It helps to classify the apps based on some parameters and suggesting the credibility of apps. using apkpure.com which is one of the famous android application markets for collecting the dataset by crawling. Web crawler which is used for indexing websites and the links related to them, and also helps in validating the hyperlinks.

A. System Architecture

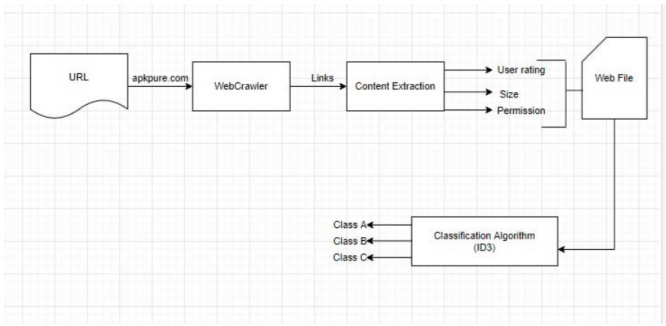


Fig. 4. The proposed system architecture

In the proposed system architecture, apkpure.com is used for collecting the dataset by crawling. The URL of the website is given as input. The use Web crawler which indexes the application and it extracts the links get content from the link and it also helps in validating hyperlinks. This is done by spider module which collects all the links from the site with filtering later and stores in two local files. The one for base links and the other one for the links from the base links. After extracting the content the three parameters size, user ratings and permissions are taken. libraries like scrappy is used for the essential feature extraction. Then it will be stored in a well-defined document. The extracted data are stored as relational database[17]. Using ID3 Decision Tree Algorithm which is an efficient algorithm should be devised to classify feature sets and finalizing the classification. A cloud is used for storing of information during classification[18].

By using ID3 decision tree algorithm, the obtained feature set is used to classify the app. For each feature classify it into class or range. Obtain the information of the entire data set. Then calculate the information of each feature. Then calculate the information gain in each feature. The one with the highest information gain is the most probable feature to classify the app. Each feature is given a range of classes after the extraction. Ratings can range from 0 to 10. Sizes are ranged as less than 20 MB more than 50 MB. Permissions can be of different aspects, different permissions depending on the importance. So it can range from 0 to 10. When a new app comes it's feature is compared with the already existing data sets of apps and classified accordingly.

VII. Web Application Attacks Detection

Web applications are permanently being exposed to attacks that exploit their vulnerabilities. the use of machine learning techniques to leverage the performance of Web Application Firewalls (WAFs), systems that are used to detect and prevent attacks[19]. A WAF is a piece of software that intercepts and inspects all the traffic between the web server and its clients, searching for attacks inside the HTTP packet contents. Once recognized, the suspicious packets may be processed in a different, secure way, for instance being logged, suppressed or derived for processing. By comparing different scenarios depending if they have valid and or attack data available for training. There are solutions for this :

- First multi-class approach for the scenario when valid and attack data is available.
- Second when only valid data is at hand.

The result of this approach that outperform MODSECURITY configured with the OWASP Core Rule Set out of the box, which is the baseline configuration setting pf widely deployed WAF technology.

Apply machine learning techniques to improve the detection capabilities of the WAF MODSECURITY giving particular importance to the task of diminishing the false positives generated by this tool when it is set up to protect a web application without reducing the true positive rate. The classification experiments have been performed on three different datasets containing labelled data.. They are PKDD (Principles and practices of knowledge Discovery in Databases), CSIC, DRUPAL.

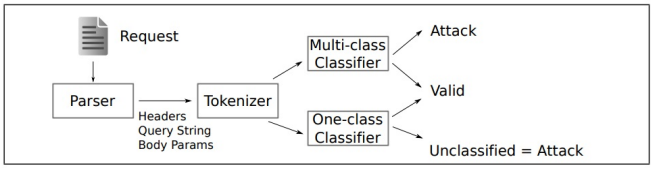


Fig. 5. Learning architecture

The parser analyses the parameters from the query string, the body and headers and split them in name and value pair as different text string The tokenization process adopted was highly dependent of the learning scenario. In sc1 we followed the classic approach of bag of words. the tokenizer performs the splitting using only spaces preserving special characters that may be included in parameters values or headers[20]. In the case of sc2 and sc3, the use of features that were blindly generated from the dataset, as in the case of sc1, did not work. The scenarios sc1 and sc2 multi-class supervised approach followed. the scenario sc3 a one class classification approach. The first approach, in which the multi-class paradigm is used, resulted in very good performance scores. The three tested classifiers provided similar results. The fact that K-NN provided good results, very close to the ones of Random Forest and SVM, indicates that if real samples of valid requests and attacks are available, the classification problem is not very complex. The second approach is also based on a multi-class paradigm, but with attack traffic collected from attacks to other applications. This approach has the advantage that only real valid traffic from the web application is needed for training. Random Forest produces more false negatives. Finally, the third approach, uses only valid requests to construct a detection model. the results outperform the ones obtained with the classic rule based MODSECURITY solution. approach reduced the number of false positives while not decreasing the true positives. consider the number of true positives generated by MODSECURITY as a reference, in two of the datasets we reduce the number of false positives (CSIC2010 and DRUPAL) and in the third one there are no improvements (PKDD2007). The results of sc1 show that machine learning algorithms can easily learn the decision boundary from training data. a. The results of sc2 show the importance of having representative attack training samples. Although a low rate of false positives was obtained, the performance on attack detection decreased. In future work we will address this issue using synthetic attacks generated with automatic tools. If only valid request are at hand, the results of the last scenario (sc3) show that a one-class solution provides many operational points that outperform MODSECURITY.

VIII. Android Malware Detection

Android Apps are freely available on Google Playstore, the official Android app store as well as third-party app stores for users to download. Due to its open source nature and popularity, malware writers are increasingly focusing on developing malicious applications for Android operating system. The effectual machine-learning based approach for Android Malware Detection making use of evolutionary Genetic algorithm for discriminatory feature selection. Android Malware analysis is of two types: Static Analysis and Dynamic Analysis. Static analysis basically involves analysing the code structure without executing it while dynamic analysis is examination of the runtime behaviour of Android Apps in constrained environment[21]. MADAM[22] provides an multi-level analysis framework where behavior of Android Apps is captured upto four levels: from package, user, application to kernel level, achieving detection accuracy as high as 96% with one disadvantage being that it could run only on rooted devices, making it incapable for commercial use. SAMADroid[23] proposed a three-way novel host server based methodology for improved performance as far as asset usage is concerned for malware detection on mobile devices. The main contribution of the work is reduction of feature dimension to less than half of original feature-set using Genetic Algorithm such that it can be fed as input to machine learning classifiers for training with reduced complexity while maintaining their accuracy in malware classification[24].

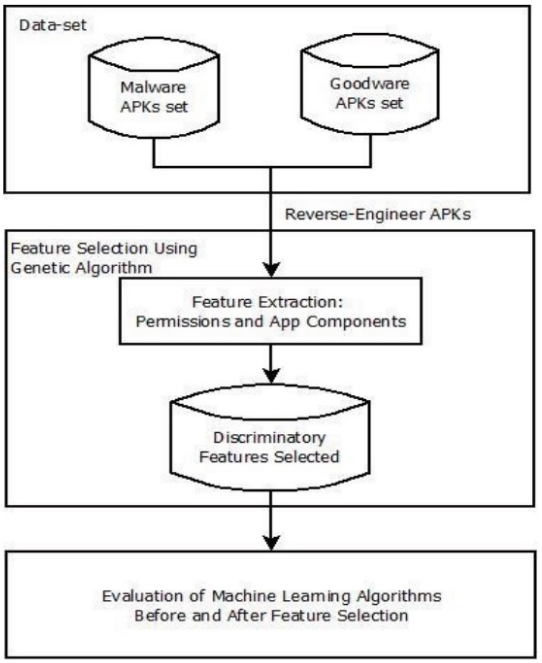

Data-set
Malware APKs set
Goodware APKs set
Reverse-Engineer APKs
Feature Selection Using Genetic Algorithm
Feature Extraction: Permissions and App Components
Discriminatory Features Selected
Evaluation of Machine Learning Algorithms Before and After Feature Selection

Fig. 6. Proposed Methodology

The figures gives a brief about proposed methodology, basically involving two units: feature extraction using Androguard tool and feature selection using Genetic Algorithm. Finally, the selected features are fed as input to machine learning algorithms for evaluation purpose. Reverse-Engineering of Android APKs, Feature Vector, Discriminatory Feature Selection where the methodologies used[24].

The steps involved in feature selection using Genetic Algorithm can be summarized as below:

1) Initialize the algorithm using feature subsets which are binary encoded such that if the feature is included it is represented by 1 and if it is excluded it is represented by 0 in the chromosome.
2) Start the algorithm defining an initial set of population generated randomly.
3) Assign a fitness score calculated by the defined fitness function for genetic algorithm.

4) Selection of Parents: Chromosomes with good fitness scores are given preference over others to produce next generation of off-springs.
5) Perform crossover and mutation operations on the selected parents with the given probability of crossover and mutation for generation of off-springs.

As the number of threats posed to Android platforms is increasing day to day, spreading mainly through malicious applications or malwares, therefore it is very important to design a framework which can detect such malwares with accurate results. The proposed methodology attempts to make use of evolutionary Genetic Algorithm to get most optimized feature subset which can be used to train machine learning algorithms in most efficient way[25].

IX. Detecting Spam Email

Machine learning models have been utilized for multiple purposes in the field of computer science from resolving a network traffic issue to detecting a malware. Emails are used regularly by many people for communication and for socialising. Security breaches that compromises customer data allows 'spammers' to spoof a compromised email address to send illegitimate (spam) emails[26]. Many tools and techniques are offered by companies in order to detect spam emails in a network. Organisations have set up filtering mechanisms to detect unsolicited emails by setting up rules and configuring the firewall settings[27]. The proposed spam detection to resolve the issue of the spam classification problem can be further experimented by feature selection or automated parameter selection for the models. This research conducts experiments involving five different machine learning models with Particle Swarm Optimization (PSO) and Genetic Algorithm (GA). This will be compared with the base models to conclude whether the proposed models have improved the performance with parameter tuning[28]. This method will experiment Bio-inspired algorithms along with Machine learning models. This will be conducted on different spam email corpora that are publicly available[29]. The main objectives are:

1) To explore machine learning algorithms for the spam detection problem.
2) To investigate the workings of the algorithms with the acquired datasets.
3) To implement the bio-inspired algorithms.
4) To test and compare the accuracy of base models with bio-inspired implementation.
5) To implement the framework using Python.

The spam detection engine should be able to take email datasets as input and with the help of text mining and optimized supervised algorithms, it should be able to classify the the email as ham or spam. The project successfully implemented models combined with bio-inspired algorithms. The spam email corpus used within the project were both numerical as well as alphabetical. Approximately 50,000 emails were tested with the proposed models. The numerical corpuses (PU), had restrictions in terms of feature extraction as the words were replaced by numbers. But the alphabetical corpuses performed better in terms of extraction of the features and predicting the outcome.

X. Detecting Wake Lock Leaks in Android Apps

The number of smartphone users worldwide is increasing; they surpass three billion and are expected to increase more[30]. The primary function of a mobile device is to connect users to the rest of the world and provide basic functionalities to aid them in performing small task. Batteries in these devices are important because they provide power to all the components of the device; they require a sufficient amount of charge to run the device for an entire day. The machine learning algorithms used to detect wake-lock leaks were first optimized using grid search to determine the best parameters. These parameters were then used in training to detect wake-lock leaks in these apps. If the energy used by an app is increasing or the app uses resources after being closed, then the app has some energy leak that requires some intervention to be resolved. Removing all resource leaks will lead to an efficient mobile app[31].

(a) Static Analysis
Static analysis allows app analysis without actually running them. We can use source code, bytecode, binary code, or intermediate representation (IR) to analyze the flow of the app.Data flow analysis is based on events, and the entry point of each event is used as the starting point of the analysis[32].

(b) Dynamic Analysis
Here flow is extracted by running the apps and performing different operations at runtime. Dynamic analysis may miss some portions of app functionality or events, as it is not possible to run all the sequences of events. To process a large number of apps, we must have multiple test cases for each app to cover the functionalities. The manual generation of these tests is time-consuming. Random test-generating tools can be used, but log files must be analyzed to extract useful information[33].

(c) Hybrid Analysis
Certain tools use a combination of the aforementioned techniques, which is known as hybrid analysis. detecting wake lock energy leaks by first extracting wake lock information and then applying a machine learning algorithm to find wake lock leaks and evaluate their accuracy[34].
Detecting wake-lock leaks is difficult, as wake locks are concentrated on a small number of methods and are called in different methods. Wake locks must be acquired and released through the flow of the program and should not cause any energy leakage. We used static analysis to address this problem and extracted the call graphs of the app for processing.

It will be first labeled the data as "clean" and leak" APKs. These APK files were preprocessed in the second section to extract useful information before they were fed into the machine learning algorithms. The preprocessed information was split into training and testing data. Training data were used to

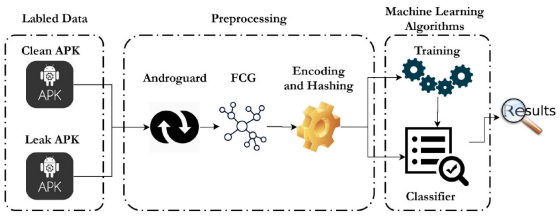


Fig. 7. Wake lock detection model using machine learning.

train machine learning algorithms to determine similarities in these data, and then a trained machine learning algorithm was used to classify the testing data. machine learning algorithms to detect wake-lock leaks and evaluate the results. The machine learning algorithms, we decompiled the Android app and constructed the FCG of wake-lock APIs, and encoded them The dataset of wake-lock leak apps was small, which led to an imbalanced data problem. . Different types of oversampling techniques were applied to balance imbalanced data problems, and it was noted that SMOTE oversampling performed well. This balanced data was then used to train different machine-learning algorithms, which were optimized using a grid search algorithm. The results indicated that SVM and SGB were the most effective and detected wake-lock leaks with high accuracy and precision[35].

XI. Conclusion

Machine learning applications are most useful applications in real life scenario. Machine learning algorithms use computational methods to "learn" information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. By using it in health sector it is helpful, and faster analysis to hand out a proper treatment plan in time are tremendously valuable. And it also detecting the malicious apps and spam mails and predicting accurate results.ML is helpful in the field of predicting information and for recommending a particular apps for the users according to their personal preferences and various other limitations. In android malware detection methodology attempts to make use of evolutionary Genetic Algorithm to get most optimized feature subset which can be used to train machine learning algorithms in most efficient way. In spam email detection successfully implemented models combined with bio-inspired algorithms machine learning algorithms to detect wake-lock leaks and evaluate the results. The results demonstrated that machine learning algorithms were effective in detecting energy leakage in mobile apps. So, machine learning is a field that is continuously being innovated, it is important to keep in mind that algorithms, methods, and approaches will continue to change.

References

[2] P. Tinschert, R. Jakob, F. Barata, J. Kramer, and T. Kowatsch, "The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps," JMIR mHealth and uHealth, vol. 5, no. 8, pp. e113, Aug. 2017.

[1] Asthma UK. (2019, Sep. 27). Asthma facts and statistics [Online]. Available: https://www.asthma.org.uk/about/media/facts-and-statistics/.

[3] My mHealth. (2020, Jan. 06). myAsthma [Online]. Available: https://mymhealth.com/myasthma

[4] J. Disantostefano, "International classification of diseases 10th revision (ICD-10)," J. Nurse Practitioners, vol. 5, no. 1, pp. 56–57, Jan. 2009.

[5] C. Oliver, "Critical realist grounded theory: A new approach for social work research," Brit. J. Social Work, vol. 42, no. 2, pp. 371–387, Mar. 2012.

[6] Shahrear Iqbal - *Spydroid A framework for employing multiple Real-time Malware Detectors on android*, 13th International Conference on Malicious and Unwanted Software (MALWARE), 2018.

[7] Zinal D.Patel: Malware Detection in Android Operating System 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2018.

[8] Dongfang Li : - Grained Android Malware Detection based on deep learning .Maejo international journal of science and technology Fine ,2018

[9] Mayank Jaiswal: Android gaming malware detection using system call analysis Indian Journal of Science and Technology, 2018

[10] Fei-Yue Wang et al. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications. IEEE Transactions on Intelligent Transportation Systems, 2010.

[11] Yongchang Ma, Mashrur Chowdhury, Mansoureh Jeihani, and Ryan Fries." Accelerated incident detection across transportation networks using vehicle kinetics and support vector machine in cooperation with infrastructure agents". IET intelligent transport systems, 4(4):328–337, 2010.

[12] Rutger Claes, Tom Holvoet, and Danny Weyns. "A decentralized approach for anticipatory vehicle routing using delegate multiagent systems". IEEE Transactions on Intelligent Transportation Systems, 12(2):364–373, 2011.

[13] C. Zhang, P. Patras, and H. Haddadi." Deep learning in mobile and wireless networking: A survey". IEEE Communications Surveys Tutorials, 21(3):2224–2287, thirdquarter 2019.

[14] Chun-Hsin Wu, Jan-Ming Ho, and D. T. Lee." Travel-time prediction with support vector regression". IEEE Transactions on Intelligent Transportation Systems, 5(4):276–281, Dec 2004.

[15] Yan-Yan Song and LU Ying. Decision tree methods: "applications for classification and prediction". Shanghai archives of psychiatry, 27(2):130, 2015.

[16] Yiming He, Mashrur Chowdhury, Yongchang Ma, and Pierluigi Pisu. Merging mobility and energy vision with hybrid electric vehicles and vehicle infrastructure integration. Energy Policy, 41:599–609, 2012.

[17] Jisha R C, Ram Krishnan, Varun Vikraman: Mobile Applications Recommendation Based on UserRating and Permission , Department of Computer Science Applications, Amrita School of Engineering, Amrita VishwaVidyapeetham, Amritapuri.

[18] Xin Su , Dafang Zhang , Wenjia Li , Wenwei Li : Android App Recommendation Approach Based on NetworkTraffic Measurement and Analysis - 20th IEEE Symposium on Computers and Communication (ISCC).

[19] A. J. Hacker and I. CISSP, "Importance of web application firewall tech nology for protecting web-based resources," ICSA Labs an Independent Verizon Business, 2008.

[20] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[21] D. Arp, M. Spreitzenbarth, M. Hübner, H. Gascon, and K. Rieck, "Drebin: Effective and Explainable Detection of Android Malware in Your Pocket," in Proceedings 2014 Network and Distributed System Security Symposium, 2014.

[22] A. Saracino, D. Sgandurra, G. Dini, and F. Martinelli, "MADAM: Effective and Efficient Behavior-based Android Malware Detection and Prevention," IEEE Trans. Dependable Secur. Comput., vol. 15, no. 1, pp. 83–97, 2018.

[23] S. Arshad, M. A. Shah, A. Wahid, A. Mehmood, H. Song, and H. Yu, "SAMADroid: A Novel 3-Level Hybrid Malware Detection Model for Android Operating System," IEEE Access, vol. 6, pp. 4321–4339, 2018.

[24] A. Firdaus, N. B. Anuar, A. Karim, M. Faizal, and A. Razak, "Discovering optimal features using static analysis and a genetic search based method for Android malware detection *," vol. 19, no. 6, pp. 712– 736, 2018.

Image Segmentation and Enhancement: An Overview

Adithyalakshmi S Kumar, Jamseena N A and Mrudula M M
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dr Reji C Joy
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: reji.c.j@vidyaacademy.ac.in)

***Abstract*—Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. Enhancement is, used to improve the visual effects and the clarity of the image. It is a type of signal processing in which input is an image and output may be an image or characteristics/features associated with that image. This paper focuses on the different image enhancement techniques. Image enhancement has been found to be one of the most important vision applications because it has the ability to enhance the visibility of images. It enhances the perceivability of poor pictures. Distinctive procedures have been proposed so far for improving the quality of digital images. To enhance picture quality image enhancement can specifically improve and limit some data presented in the input picture. To enhance images, have to go through some image enhancement techniques which remove the noises in the images. For that noise removal, there are some segmentation methods also discussed.**

***Index Terms*—Image processing, Image enhancement, Contrast enhancement, Histogram Equalization, Texture, Segmentation, Noise Removal, Spatial Domain, Frequency Domain, Contrast Limited Adaptive Histogram Equalization, Multiscale morphology, Region Of Interest, Image compression.**

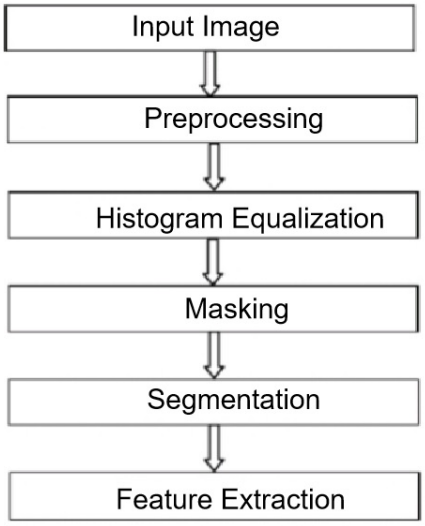


Fig. 1. Flowchart for segmentation process

I. Introduction

Image enhancement is essential for any image processing application. The objective of image enhancement is to reveal the hidden information which is not available for the purview of the observer due to the presence of low and poor contrast during image acquisition. The quality of the image can be raised by increasing the contrast. Contrast enhancement has found a prominent application in various fields such as medical, and satellite imaging systems owing to its better visibility of the features. In this paper, a comprehensive survey on different contrast enhancement techniques exclusively on spatial domain and frequency domain is presented. Image processing is a method for converting an image to a digital image or aspect and performing certain functions on it, to get an enhanced image or useful information from it. One has to undergo three phases generally through a data for getting a digital image and the phases are: pre-processing, enhancement and display, and information extraction. Contrast enhancement can be defined as changing the input image's pixels intensity to utilize the maximum possible bins. An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of colour or intensities in an image or selected region of an image.

II. Image Segmentation

Image segmentation is a method in which a digital image is broken down into various subgroups called Image. segments which help in reducing the complexity of the image to make further processing or analysis of the image simpler. Segmentation in easy words is assigning labels to pixels. All picture elements or pixels belonging to the same category have a common label assigned to them. see figure:1

1) Fibroid Ultrasound Image: Reads Original Fibroid contained Ultrasound Image.
2) Preprocessing: Preprocessing is a process to remove the noise, as Ultrasound Image contains speckle noise, so

the images are not clearly visible (low contrast), and it's difficult to detect the edges. Hence, preprocessing must be done to clean the image, there are various noise removal techniques that have advantages and also disadvantages.

3) Histogram Equalization: it means distributing the intensity values or pixel values uniformly throughout the image. Most of the time when an image contains noise or brightness variation, the intensity values are not uniformly distributed. When intensities are adjusted, the brightness level is high at peak areas and is compressed at valleys. The greyscale pixel values are uniformly distributed from 0 to 255.
4) Masking: A mask is also known as a filter or spatial filtering. The filtering process is nothing but a mask convolution with an image. Masks of various sizes can be created such as 3x3, 5x5, or 7x7, and the mask moves over an original image. Masks are used to increase the sharpness of an image and to detect the edge details.
5) Segmentation: Segmentation is nothing but decomposing the image into its constituent parts or regions. These regions are classified as normal and abnormal regions. An abnormal region contains unwanted masses or tumors, those parts must be segmented. Segmentation may fail or succeed, it depends upon the accurate detection of the tumor and exact segmentation.
6) Feature Extraction: The Feature Extraction is nothing but extracting the features that are required to detect the abnormalities from the 2D image.
   There are some Image Segmentation Techniques:-
   - Edge Based Segmentation
   - Region-Based Segmentation
   - Clustering Based Segmentation

A. Texture Enhancement and Segmentation

Image segmentation is an important research direction in pattern recognition and image understanding, but existing texture segmentation algorithms cannot take full advantage of some texture information of texture image, such as the direction, width, density of ridge line, and so on, and can also not effectively carry out the segmentation of various texture image quality. In order to efficiently implement the texture image segmentation, strengthen the amassing of region segmentation, improve the accuracy of segmentation, and achieve more accurate target recognition, this paper defines the direction of the texture, calculates the width of the ridge line, gives the distance characteristics between textures, and establishes the mathematical model of the texture border, accordingly presents a new texture segmentation algorithm and compares with other texture segmentation algorithms.

III. Image Enhancement Techniques

Image enhancement is essential for any image processing application. The objective of image enhancement is to reveal the hidden information which is not available for the purview of the observer due to the presence of low and poor contrast during image acquisition. There are basically two types of images enhancement techniques,

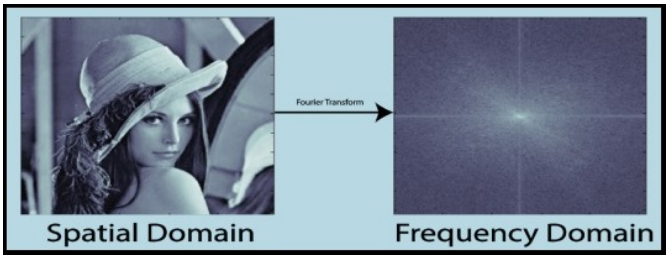


Fig. 2. fourier transformation

A. Spatial Domain Technique

Spatial domain enhancement techniques deal with the images direct intensity values. The spatial domain is used to define the actual spatial coordinates of pixels within an image. Spatial domain techniques like the logarithmic transforms, power-law transform, and histogram equalization, are based on the direct manipulation of the pixels in the image. Spatial techniques are particularly useful for directly altering the gray level values of individual pixels and hence the overall contrast of the entire image. But they usually enhance the whole image in a uniform manner which in many cases produces undesirable results.

B. Frequency Domain Technique

In the frequency domain, a digital image is converted from the spatial domain to the frequency domain. In the frequency domain, image filtering is used for image enhancement for a specific application. A Fast Fourier transformation is a tool of the frequency domain used to convert the spatial domain to the frequency domain. For smoothing and image, the low filter is implemented and for sharpening an image, a high pass filter is implemented. Edges and sharp transitions (e.g. Noise) in an image donate meaningfully to the high-frequency content of the Fourier transform. Low-frequency contents in the Fourier transform are responsible for the common appearance of the image over smooth areas. see fig2

IV. Contrast Enhancement Techniques

Contrast enhancement has found a prominent application in various fields such as medical, and satellite imaging systems owing to its better visibility of the features. This paper is a comprehensive survey on different contrast enhancement techniques.

- Histogram Equalization: In HE the dynamic ranges of the histogram of the input image are increased and produce uniformly distributed intensities in a histogram as an output which implies a good quality enhanced image. see difference in images in fig 3
- Bi-Histogram Equalization : This technique is used to control the brightness of images by dividing the input image's histogram into sub histograms based on the

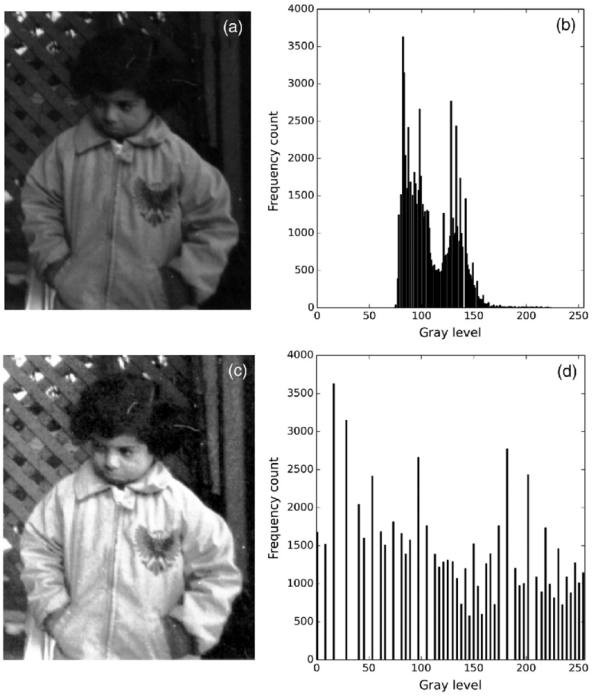


Fig. 3. Non-equalized and Equalized images and their correspondent histograms

Fig. 4. Difference between images that clahe applied and not

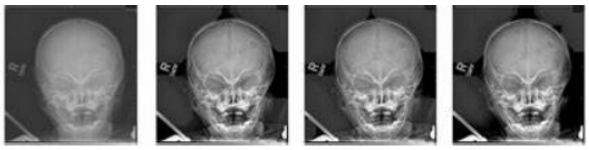


Fig. 5. morphological enhancement on a medical image of a skull

mean value and then they undergo equalization separately resulting in a flatter histogram.

- Gray level grouping : According to Zhiyu Chen the researcher, made an automatic method for contrast enhancement by implementing Gray-Level-Grouping in which the first group of the histogram components of a low-contrast image into a proper number of bins according to a selected criterion, then redistribute these bins uniformly over the grayscale, and finally ungroup the previously grouped gray-levels.
- Dynamic Histogram Equalization: In this method, HE can be done without any loss in the image, here the image is divided into sub-images according to their local maxima, then apply dynamic histogram equalization to every sub-image and then combine those sub-images.
- Adaptive Histogram Equalization : This technique is not like normal histogram equalization because it transforms each pixel with a transformation function derived from a neighbourhood region.
- CLAHE (Contrast Limited Adaptive Histogram Equalization): enhances local contrast of an image without an increase in the noise of the image. It is based on HE and can say as an extension of HE. Here tiles are defined as different sections of a full image. These tiles are enhanced to redistribute the pixel values. And then tiles are combined using bilinear interpolation. see fig 4
- Top Hat Transformation: This is a tool used for extracting bright or dark features. There are white top hats used to extract the white or bright portion and a Blacktop hat transformation used to extract the dark or black features of the image. In a white top hat, Subtracting the opened image from the original one produces an image where the features that have been removed by opening are clearly visible. In closing operation instead of an opening and subtracting the original image from the closed one helps to extract dark features from a brighter background. This is in the case of a black top hat which is opposite to the white top-hat transformation.
- Morphological Enhancement: here object in the image is considered an A set. Operations are defined between object and structuring elements. Morphological Filtering on a single scale: This technique is a local contrast enhancement technique by using Top Hat Transformation based on structuring elements. This technique is used in medical images like digital mammograms by using one certain output image as the next input image to get a clear picture. Morphological filtering on Multiscale: In this technique, the extracting features is done based on size and shape. This technique is used in vascular angiograms where arteries are checked and several medical images. morphological enhancement on a skull is done in fig 5

V. Image Compression Based on Region of Interest

The problem of using regions of interest(ROI) to increase the efficiency of image/video transmission over constrained-bandwidth channels has received considerable attention in the image processing literature. This is due to the fact that various applications can strongly benefit from the uneven distribution of resources(bits, error protection, resolution, etc) through different image regions. The main target of the region of

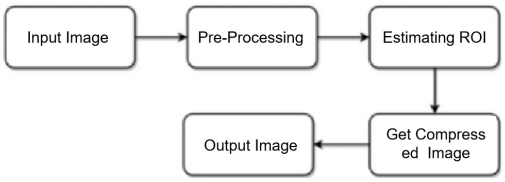


Fig. 6. flowchart of image compression

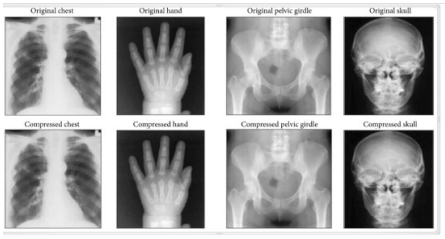


Fig. 7. Compression can method for medical images based on ROI

interest (ROI) based compression for the medical image is to improve the compression efficiency for transmission and storage. The main goal of Region of interest(ROI) compression is to compress ROI with main quality as compared to other regions call background. For example, while compressing a medical image the system's important region should be compressed with better quality than the background. Thus, the ROI area is compressed with less compression ratio and the background with the highest possible compression ratio in order to get overall better compression performance. The difference between compressed and uncompressed images for medical based on ROI is seen in fig 7

Image compression can improve the quality of the images. There are two types of image compression which are lossless and lossy.

- Lossless compression: there will be no loss in the image after compression. The goal of lossless image compression is to represent an image signal with the smallest possible number of bits without loss of any information, thereby speeding up transmission and minimizing storage requirements. see figh to see a lossless compressed image.

- Lossy compression: when an image is compressed some part of the image will be completely lost. Mainly lossy compressions are done in jpeg. Lossy image compression is a process that removes some of the data from your image file, reducing the overall file size. This process is irreversible, meaning that the file information will be removed permanently. see fig 9 to know the difference.

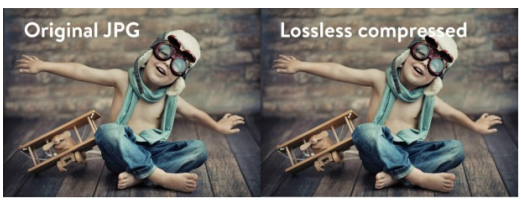


Fig. 8. lossless compressed images

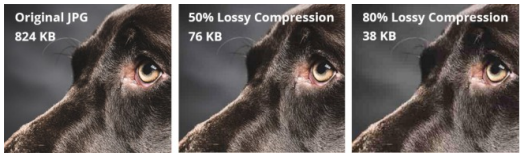


Fig. 9. lossy compression

Fig. 10. underwater image enhanced

VI. APPLICATIONS

A. Image Enhancement Techniques for Underwater

Underwater images mainly suffer from the problem of poor color contrast and poor visibility. These problems occurred due to the scattering of light and refraction of light while entering from a rarer to a denser medium. Scattering causes the blurring of light and reduces the color contrast. These effects of water on underwater images are only not due to the nature of water but also because of the organisms and other material present in the water. To overcome these issues, the acquired images are adjusted using color balance and then corrected using a white balance. Many techniques and methods are established by researchers to solve the problem of underwater image enhancement.see fig 10

B. Image Enhancement Techniques for Medical Images

Medical imaging is the technique and process of imaging the interior of a body for clinical analysis and medical intervention, as well as visual representation of the function of some organs or tissues. Medical imaging seeks to reveal internal structures hidden by the skin and bones, as well as to diagnose and treat disease. Edge detection plays an important role in medical imaging because all information has preserved in edges. Fuzzy Histogram Hyperbolization is one of the techniques in medical images. FHH is found to be applied successfully to the area of image enhancement. Moreover, FHH can be applied not only for image enhancement but also through the selection of a suitable membership function, it can be used in other fields of image processing. see fig 11

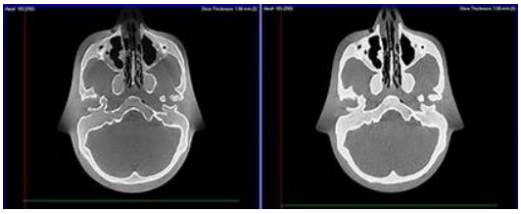


Fig. 11. dental image enhanced

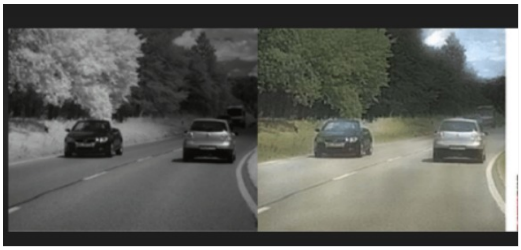


Fig. 12. colorization for infrared image

C. Image Enhancement Techniques of Infrared Images

Capturing of infrared images is an easy task but perceptual visualization is difficult due to environmental conditions such as light rain, partly cloudy, mostly cloudy, haze, poor lightning conditions, noise generated by the sensors, geographical distance, and appearances of the objects. To improve the human perception and quality of the infrared images for further processing like image analysis, image enhancement is an essential process. see fig 12

VII. Conclusion

Image Enhancement is one of the most important and difficult techniques in image research. The aim of image enhancement is to improve the visual appearance of an image or to provide a better transform representation for future automated image processing. In enhancement techniques, morphological filtering is better than other techniques because it implements mathematical morphology in images. The size of the images can be reduced by using compression techniques. They are loss and lossless. Most common we are using loss compression why because in all system applications which we generally use will support jpg file formats. The aim of image enhancement is to improve the visual appearance of an image or to provide a "better transform representation for future automated image processing. Many images like medical images, Satellite images, aerial images, and even real-life photographs suffer from poor contrast and noise. It is necessary to enhance the contrast and remove the noise to increase image quality.

References

[1] M.M. Raheman, P.K. Mithun Kumar, B. Borucki, K.S. Nowińskiand and M.S. Uddin, "Speckle noise reduction of ultrasound images using Extra-

[2] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Martin-Fernandez, C. Alberola-López and S. Aja-Fernández, "Anisotropic diffusion filter with memory based on speckle statistics for ultrasound Images", IEEE Transactions on Image Processing, Vol. 24, No. 1, pp. 345-358, 2015.

[3] K.M. Prabusankarlal, P. Thirumoorthy, and R. Manavalan, "Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound", Human-centric Computing and Information Sciences, Vol. 5, No. 1, Pp. 1-17, 2015.

[4] N. Yahya, N.S. Kamel and A.S. Malik, "Subspace-based technique for speckle noise reduction in ultrasound images", Biomedical engineering online, Vol. 13, No. 1, Pp. 1, 2014.

[5] H.K. Gill, A. Girdhar and J. Kaur, "Automatic region growing segmentation for medical ultrasound images", In International Conference on Medical Imaging, m-Health, and Emerging Communication Systems (MedCom), Pp. 454-457, 2014.

[6] J. Saranya and S. Malarkhodi, "Filtering and segmentation of a uterine fibroid with ultrasound images", In 2012 International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME), Pp. 296-301, 2012.

[7] N.H. Mahmood, N. Zulkarnain and N.S.A. Zulkifli, "Ultrasound liver image enhancement using watershed segmentation method", (IJERA), Vol. 2, No. 3, Pp. 691-694, 2012.

[8] N.M. Zayed, A.M. Badwi, A. Elsayad, M.S. Elsherif and A.M. Youssef, "Wavelet segmentation for fetal ultrasound images", In Circuits and Systems. Proceedings of the 44th IEEE 2001 Midwest Symposium on, Vol. 1, Pp. 501-504, 2001.

[9] Haris Acker, mohd.A.Saleh, Ali abd misreb "A Review On Image Enhancement Technique " Article in Southeast Europe Journal of Soft Computing · April 2019

[10] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active contour models", International journal of computer vision, Vol. 1, No. 4, Pp. 321-331, 1988.

[11] S. Sridevi and M. Sundaresan, "Survey of image segmentation algorithms on ultrasound medical images", In International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), Pp. 215-220, 2013 .

[12] Rashmi Choudhary, Sushopti Gawade Department of Computer Engineering PIIT, Mumbai University, India. "Survey on Image Contrast Enhancement Techniques"

[13] M.Joseph Prakash1 , Dr.V.Vijaya Kumar2 & Dr.A.VinayaBabu3 1Research Scholar, JNTUK University, Kakinada ,2Department of CSE, GIET, JNTUK University, Rajahmundry, 3Department of CSE, JNTUHCE, JNTUH University, Hyderabad, "MORPHOLOGY BASED TECHNIQUE FOR TEXTURE ENHANCEMENT AND SEGMENTATION"

[14] Rupneet Kaur Hanspal1 , Kishor Sahoo2 1Central College of Engineering and Management, Kabir Nagar, Raipur, Chhattisgarh Swami Vivekananda Technical University, Bhilai 2Central College of Engineering and Management, Kabir Nagar, Raipur, HOD, Electronics and Telecommunication, Chhattisgarh Swami Vivekananda Technical University, Bhilai on "A Survey of Image Enhancement Techniques"

[15] Ritika Assistant Professor Guru Nanak Dev University Regional Campus, Gurdaspur and Sandeep Kaur Assistant Professor Guru Nanak Dev University Regional Campus, Gurdaspur on "Contrast Enhancement Techniques for Images– A Visual Analysis"

[16] Dawood, F. A., Rahmat, R. W., Kadiman, S. B., Abdullah, L. N., &Zamrin, M. D. Measurements of mitral annular displacement in 2D echocardiography images, American Journal of Applied Sciences, 295-303, vol 12, issue 5, 2015.

[17] Amiri, S. A., &Hassanpour, H. A preprocessing approach for image analysis using gamma correction. Department of Computer Engineering Shahrood University of Technology Iran. 2012

[18] Min, B. S., Lim, D. K., Kim, S. J., & Lee, J. H. A novel method of determining parameters of CLAHE based on image entropy. 2013

[19] Rafael.C.Gonzalez and Richard.E.Woods "Digital image processing" Third edition.

[20] Ziad M. Abood, Zahra S. Dheef&Lamyaa M. Kadhim, 2017, Diagnosis of colon tumor tissue using fuzzy logic Vol. 12, No.1, March - IJST.

[21] Sargun and Sashi Rana "Medical Image enhancement techniques for Image Processing" Dept of E&T, International journal of current engineering and technology,Vol 5, No.2(April 2015).

[22] Hasan Demirel, Cagri Ozcinar, and Gholamreza Anbarjafari, "Satellite image contrast enhancement using discrete wavelet transform and singular value decomposition", IEEE geoscience and remote sensing letters, vol 7, issue 2, April 2010.

E-Commerce Security

Athira K M, Haritha H and Manasa K Mani
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Dijesh P
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: dijesh.p@vidyaacademy.ac.in)

***Abstract*—The term electronic commerce (E-commerce) refers to a business model that allows companies and individuals to buy and sell goods and services over the Internet. E-commerce increases efficiency and flexibility of trading. E-commerce Security is a part of Information Security framework. It includes Computer Security, Data security and other wider realms of Information Security framework. It is the set of guidelines that are designed to allow safe transactions on the web. Privacy and security are considered to be the major factor in electronic technologies. Now a days Consumer privacy is becoming the most publicized security issue replacing theft and fraud as top concerns in e-commerce. User and system administrator awareness is becoming more important in the effort to counter e-commerce attacks. Consumers are slowly becoming aware of some security features such as encryption. Consumer disposition to the information quality of the website, trust, privacy concerns, reputation, security concerns, and the company's reputation have strong effects on Internet consumers' trust in the website. Information security, therefore, is an essential management and technical requirement for any efficient and effective payment transaction activities over the internet. E-commerce security is also the protection of e-commerce assets from unauthorized access, destruction, alteration, or use so its dimensions to be studied are- Integrity, Privacy, Non-repudiation, Authenticity, Confidentiality, and Availability. During various phases of an electronic transaction, the information such as product specification, order details, payment, and delivery information travels over the comparison of the two most popular E-commerce transactions secure protocols SSL and SET. Peer-to-Peer, commonly abbreviated to P2P, is any distributed network architecture composed of participants that make a portion of their resources directly available to other network participants, without the need for central coordination instances. P2P applications include key characteristics of anonymity, autonomy, and decentralization. As the technical advantages of P2P, e-commerce based on P2P technology is now in its stage of development, and many people are dedicated to the study. P2P e-commerce as a new commercial mode, there are still important challenges to overcome before the full potential of P2P e-commerce can be really realized. In this paper, we discuss some security related issues about e-commerce and P2P e-commerce.**

***Index Terms*—E-commerce, Security, Privacy, Trust Peer to peer e-commerce, Security protocols**

I. Introduction

E-commerce Security is a part of the Information Security framework and is specifically applied to the components that affect e-commerce that include Computer Security, Data security and other wider realms of the Information Security framework. E-commerce security has its own particular nuances and is one of the highest visible security components that affect the end user through their daily payment interaction with business. Security in E-commerce is very important part since communication can be easily intercepted, messages can be inserted, and the absolute identity of involved parties may be uncertain. Security tries to accomplish tasks such as authentication, data integrity, confidentiality and non-repudiation. E-commerce based on P2P technology is now in its stage of development, and many people are dedicated to the study. Peer-to-Peer, commonly abbreviated to P2P, is any distributed network architecture composed of participants that make a portion of their resources directly available to other network participants, without the need for central coordination instances.P2P e-commerce as a new commercial mode, there are still important challenges to overcome before the full

Fig. 1. Security of E-Commerce

potential of P2P e-commerce can be really realized.

1) Impact on marketers and retailers According to economists, e-commerce has led to increase in price competition because it has increased user's ability to collect information regarding product and prices. Its growth has also affected industry structure in two areas that has seen growth in e-commerce, bookshops and travelling agencies.
2) Impact on supply chain management For a long time, companies had been troubled by gap between benefits which supply chain technology had and solutions to deliver those benefits.
3) Social impact Along within e-commerce and its unique charm that had appeared gradually, virtual enterprise, virtual bank, network marketing, online shopping, payment and advertising, such this new vocabulary which is unheard-of and now had become as familiar to people

II. Types of E-commerce

There are three types of E-commerce. They are

- B2B E-Commerce: It involves online order transactions typically between wholesales, manufacturers, and distributors. Pricing based on the quantity.
- B2C E-Commerce: It is used to describe a commerce transaction between a business and an end consumer. This transaction is done by a catalogue utilizing shopping cart software.
- C2C E-Commerce: It refers to the sale of goods or services from one consumer to another. PayPal is an example.

III. Advantages and Disadvantages

A. Advantages

- With the help of e-commerce, organizations can expand their market to worldwide markets with minimum investment.
- Customers can enquire about a product or service and place orders anytime, anywhere from any location.
- E-commerce helps in reducing the cost of products and organizations to reduce the cost and manage the paper-based information by digitizing the information.
- E-commerce improves and manages the brand image of the company.
- E-commerce helps organization to provide better customer services.
- E-commerce helps the government to deliver public services such as healthcare, education, social services at a reduced cost and in an improved manner.
- E-commerce increases the productivity of organizations.
- In online shopping there is facility of writing feedback of product, with the help of this feedback one can easily understand the quality of product or any drawbacks about the product
- E-commerce provides users with multiple options to compare and select the cheaper and better options.
- A customer can see the accurate detailed information within fraction of seconds, rather than waiting for days.
- E-Commerce increases the competition among organizations so organizations provide bumper discounts to customers.
- Customers need not travel to shop a particular product, thus there is less traffic on road and also less air pollution.
- E-commerce also helps in rural areas to access services; these services otherwise are otherwise not available to them.

B. Disadvantages

- E-commerce biggest disadvantage is lack of system security
- Sometimes network bandwidth might cause an issue in some places.
- Special types of web servers or other software might be required by the vendor.
- Sometimes, it becomes difficult to integrate an e-commerce software or website with existing applications or databases.
- Chances of software/hardware compatibility issues.
- The cost of creating e-commerce application in-house is very high.
- It is difficult to trust the customers to security or privacy on online transactions.
- Lack of touch or feel of products during online shopping is one of the biggest drawbacks.
- Internet access is still not cheaper and is inconvenient to use for many customers, those are living in remote villages.

Privacy, security and trust are major concerns in e-commerce. Privacy is the control over one's personal data whereas and security is the attempted access to data by unauthorized users. There are four categories of antecedents that influence consumer trust and consumers' perceived risk towards electronic commerce entities which are:

1) Experience-based: e.g., e-commerce experience, familiarity, Internet experience, etc.
2) Cognition (observation)-based: e.g., system reliability, privacy protection, quality of information, security protection, brand image, etc
3) Personality-oriented: e.g., disposition to shopping habits, trust, etc
4) Affect-based: e.g., presence of third-party seals, reputation referral, recommendation, buyer's word-of-mouth, feedback, review comments, etc

IV. Privacy in E-commerce

Consumers have two kinds of privacy concerns.

- They are concerned about the risk of secondary use i.e., the reuse of their personal data for unrelated purposes without their consent such as sharing with third parties who were not part of the transaction in which the consumer related his or her personal data.

- consumers are concerned over unauthorized access to personal data because of security breaches or the lack of internal controls.

V. TRUST IN E-COMMERCE

Trust is an important issue in e-commerce, because unlike real world transactions, the retailer is not present in person during the transaction and the consumer is not dealing with a real person. Consumer is dealing with the interface. It is also much more difficult for customers to determine the authenticity of websites and also difficult to connect identities with actual individuals. For consumers, security, privacy, functionality and user-friendliness issues are considered to be barriers to online shopping. Apart from that user likes to keep their personal information private. The probable risk is higher in e-commerce mainly because of unawareness, proximity and negligible physical interactions. So, in order to see why consumer, engage or do not engage in e-commerce, it is important to study their online trust in e-commerce as a marketplace. Some factors which are the most important to the consumers for trusting a website are the security systems, privacy, reputation of the company, payment methods offered by it, customer service provided, the website design, control of technology, ease of usage, user friendliness of the website, and the price offered by the company. It is also important to consider the age group of people who uses internet as market place.

VI. SECURITY

E-commerce Security is a part of Information Security framework and is specifically applied to components that affect e-commerce that include Computer Security, Data security and other wider realms of Information Security framework. Fig 1 Security of E-Commerce The safety of e-commerce had its actual degree and is one of highest visible security components that affect end user through their daily payment interaction within business. Today, privacy and security are a major concern for electronic technologies.

E-commerce security is the protection of e-commerce assets from unauthorized access, use, alteration, or Destruction. While security features do not guarantee a secure system, they are necessary to build a secure system.

- Authentication: It enforces that you are the only one allowed to logon to your Internet banking account.
- Authorization: Allows only you to manipulate your resources in specific ways.
- Encryption: Deals with information hiding.
- Auditing: Keeps a record of operations. Merchants use auditing to prove that you bought specific merchandise.
- Integrity: prevention against unauthorized data modification
- Non repudiation: prevention against any one party from reneging on an agreement after the fact.
- Availability: prevention against data delays or removal.

A. Security Threats

Three types of security threats exist they are:

1) Denial of service
2) Unauthorized access
3) Theft and fraud

1) Denial of Service (DOS): Two major types of DOS attacks: Spamming and viruses.

- Spamming - is sending spontaneous business messages to people, E-mail bombing caused by a hacker targeting one computer or network, and sending a huge number of email messages to it.
- DDOS (Distributed denial of service) - Includes programmers putting programming specialists onto various outsider frameworks and setting them off to at the same time send solicitations to a proposed target.
- Viruses - self-reproducing PC programs intended to perform undesirable occasions.
- Worms - Exceptional infections that spread utilizing direct Internet associations.
- Trojan Horses - It is one amongst the worst network security threats where attackers use these programs to swipe sensitive information from their computers with ease.

2) Unauthorized access : The illegal access to frameworks, systems, applications, information or data, Passive unauthorized access- listening to communications channel for discovering privileged insights. It might utilize content for destructive purposes, Changes intent of messages, e.g., to prematurely end or postpone a negotiation on a transaction. Active unauthorized access-modifying and altering or monitoring system or data. Message stream modification and changing the messages.

3) Theft or Fraud: It occurs when the stolen information is utilized or adjusted. Theft of software through illegal replicating from organization's servers. Robbery of equipment, specifically laptops or workstations is also among the major security threat.

B. Attacks

There are two attacks E-commerce. There are:

1) Passive Attacks: In passive attacks, the attacker observes the messages, then copy and save them and can use it for malicious purposes. They can be dangerous for confidentiality of the message.
2) Active Attacks: In active attacks, the attacker intercepts the connection and efforts to modify the message's content. It is dangerous for integrity and availability of the message. Active attacks involve Modification of message, Repudiation, and Denial of service.

C. E-commerce Security Tools

There are various tools are used to provide the security:

1) Digital Signatures: These are the public-key primitives of message authentication. In the physical world, it is

common to use handwritten signatures on handwritten or typed messages. They are used to bind signatory to the message. Similarly, a digital signature is a technique that binds a person/entity to the digital data.

2) Encryption software and techniques: It is a software that can encrypt and decrypt data, often in the form of files on a hard drive or packets sent over a network. It is a fundamental part of modern computer communications and file protection. There are various encryption software available in the present market. For e.g., Bit locker, VeraCrypt, Axcrypt.
3) Public Key infrastructure: Public key infrastructure is a set of roles, policies and procedures needed to create manage, distribute use public key encryption.
4) Biometrics: It refers to authentication techniques that rely on measurable physical characteristics that can be automatically checked. There are several types of biometric identification schemes – face, fingerprint, retina, signature, iris, etc.
5) Firewalls: These are frequently used to prevent unauthorized internet users from accessing private networks connected to the internet, especially intranets. Firewalls can be either hardware or software but the ideal firewall configuration will consist of both. A firewall acts as a barrier between a trusted network and an untrusted network. A firewall controls access to the resources of a network through a positive control model.
6) Algorithm: Algorithms are also used for encrypting the data, there are various algorithm to change the plaintext into chipper text. Algorithm: RC4, CRT

D. Secure E-Shopping Guidelines and Solution

1) Shop at Secure Web Sites: How can we check whether a Web site is secure or not? Secure websites utilize encryption techniques to exchange data or information from your computer to the online trader's website. Here's the manner by which you can tell when you are working with a secure website:
   - If you look at the address bar of your browser where the Web site address is shown, you are supposed to see https://. The "s" that is shown after "http" demonstrates that the Website is secure.
   - Another approach to find out whether a Web site is secure is to search for a shut lock displayed on the address bar of the website. If that lock is modifiable, you should presume it is not a safe website.
2) Research the Web Site before You Order: Work with organizations you already know. In the event that the organization is new, do your homework before purchasing their items. If you decide to buy something from an obscure organization, start with a low-priced order to learn if the organization is reliable. Dependable organizations ought to advertise their physical business address and at least one telephone number, either customer service or a helpline.
3) Read the Web Site's Privacy and Security Policies: Every trustworthy online Website provides data about the manner it processes your request. It is typically recorded in the area entitled —Privacy Policy. You can check out whether the vendor expects to share your data with a third party or associate organizations.
4) Be Aware of Cookies and Behavioural Marketing: Online vendors and in addition different locales watch our shopping and surfing propensities by utilizing "treats," a web based following framework that connects bits of code to our Internet programs to track which destinations we visit as we pursuit the Web.
5) What's Safest: Credit Cards, Debit Cards, Cash, or Checks. The most secure approach to shop on the Internet is with a Visa. In the occasion something turns out badly, you are ensured under the government Fair Credit Billing Act. You have the privilege to question charges on your Visa, and you can withhold instalments amid a lender examination. Similarly, as with checks, a platinum card uncovered your financial balance to cheats. Your financial records could be wiped out in minutes.
6) Never Give Out Your Social Security Number: Giving your Social Security number is not a necessity for putting in a request at a web-based shopping website. There is no requirement for the vendor to request it. Giving out your Social Security number could prompt have your character stolen.
7) Keep Your Password Private: Numerous web-based shopping destinations require the customer to sign in before putting in or seeing a request. The customer is normally required to give a username and a secret word. Never uncover your secret word to anybody. Try not to reuse a similar secret key for different locales and the best secret key has no less than eight characters and incorporates numbers and letters.
8) Don't Fall for "Phishing" Message: Identity thieves send massive numbers of emails to Internet users that ask them to update the account information for their banks, credit cards, online payment service, or popular shopping sites. The email may state that your account information has expired, been compromised or lost and that you need to immediately resend it to the company.

E. E-commerce Security Protocols

1) Secure Sockets Layer (SSL) SSL is a security protocol protects communications between any SSL-enabled client and server software running on a network that uses TCP/IP, Gopher, FTP, Telnet etc. SSL approach is to add a layer on top of the existing network transport protocol and beneath the application. This approach applied by adding an intermediate step, requiring negotiation of secure transmission options, to the establishment of a network connection. During SSL connection establishment only, the server is authenticated using a digital certificate. SSL steps to establish a secured connection between the customer client and server.

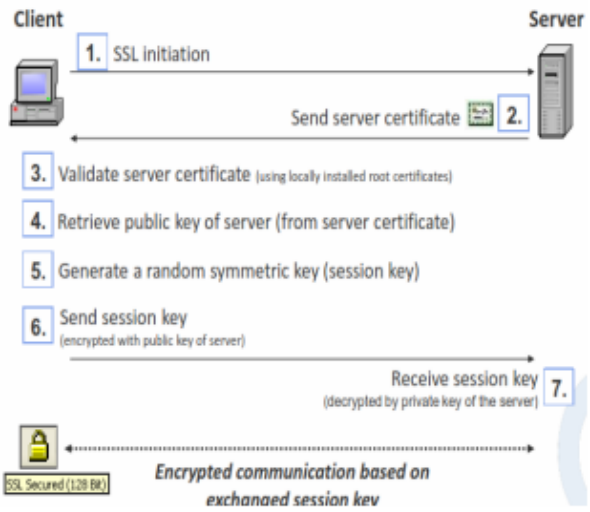

Client
Server
1. SSL initiation
Send server certificate 2.
3. Validate server certificate (using locally installed root certificates)
4. Retrieve public key of server (from server certificate)
5. Generate a random symmetric key (session key)
6. Send session key
(encrypted with public key of server)
Receive session key 7.
(decrypted by private key of the server)
SSL Secured (128 Bit)
Encrypted communication based on
exchanged session key

Fig. 2. SSL Secured connection steps

Advantages of SSL are: Transparency, Ease of use for customer, Low complexity.
Disadvantages of SSL are: The merchant cannot reliably identify the card holder, SSL only protects the communication link between the customer and the merchant, without a third-party server, SSL cannot provide assurance of non-repudiation.

2) **Secure Electronic Transactions (Set)**
It is a standardized industry wide protocol specification designated to secure payment transactions and authenticate the parties involved in the transaction in any type of networks including Internet. SET was created to provide the trust needed for consumers. The protocol uses cryptography and digital certificates to provide confidentiality of the information, ensure payment integrity, and authenticate merchants. SET Specifications:

- SET uses RSA cryptosystem for encrypt and decrypt the transaction packets along with the use of digital certificates and digital signature for authentication of all parties to the transaction and validation.
- It makes the online transaction safer by using the digital certificates to verify that consumers and merchants are both authorized to use and accept Visa cards.
- SET incorporates the use of public key cryptography to protect the privacy of personal and financial information.
- With SET, cardholders can validate that the Internet merchant is legitimate through the merchant's digital certificate. SET software automatically checks that merchant has a valid certificate representing their relationship with their financial institution. This provides consumers with the confidence that their payments will be handled with the same Visa promise that they trust today.

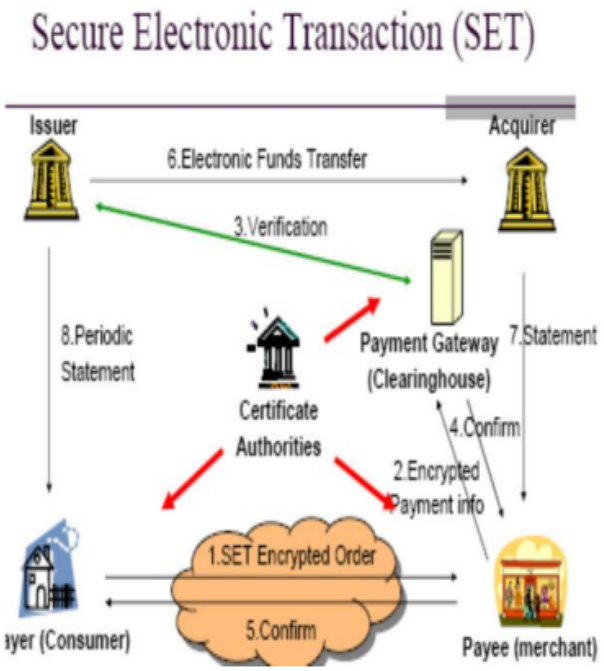


Fig. 3. Secure Electronic Transaction(SET)

Advantages of SET protocols are: Confidentiality, authentication and data integrity was verified by a large collection of security proofs based on formal methods, In the standard variant of the protocol, SET prevents merchants from seeing the customer payment information, since this information is encrypted using the payment gateway's public key and to ensure merchant privacy.
Disadvantages of SET are: The customer must install additional software, which can handle SET transactions. The customer must have a valid digital certificate. Implementing SET is more costly than SSL for merchants as well. Adapting their systems to work with SET is more complicated than adapting them to work with SSL. SET employs complex cryptographic mechanisms that may have an impact on the transaction speed

VII. P2P E-COMMERCE

Peer-to-Peer, commonly abbreviated to P2P, is any distributed network architecture composed of participants that make a portion of their resources directly available to other network participants, without the need for central coordination instances. The goal of P2P e-commerce is to present a long perspective on the evolutionary a new, fine grained type of e-commerce on the Internet.

A. Characteristics of P2P application

- A high degree of autonomy: which means that the interactions between peers, the activities of resources sharing, and the activities of processing distributed data are independent without any authorization from a centralized server.
- Anonymity : Anonymity means nodes or participants are anonymous or pseudonymous.
- No fixed centralized nodes: P2P networks have no fixed central nodes, and all participants compose the distributed network architecture.
- A high degree of capacity: P2P breaks the bottleneck of server and bandwidth, the more nodes, the better transmission speed of network traffic.

B. P2P E-commerce Security Requirements

- Authentication: It involves confirming the identity of a person .Which includes digital signature, watermark, shared secret, public key infrastructure, and so on.
- Non-repudiation: It ensures a participant in a transaction cannot repudiate, or refute the validity of a statement or contract. Most common non-repudiation method through is data signature.
- Integrity: It has to do with perceived consistency of actions, values, methods, measures, principles, expectations and outcome. To ensure integrity uses the technology of digital signature.
- Authorization: The combination of authentication and authorization could protect the interest of consumers in maximum.
- Confidentiality: Some sensitive data should be protected strictly, encryption is the best way to keep the confidentiality.
- Availability: To ensure that a legitimate user can access the information properly, it is necessary to ensure the availability of P2P e-commerce.
- Controllability: In P2P e-commerce there lot of malicious peers who bring security attacks, so controllability play an important role.

C. Security Problems in P2P E-commerce

- The problem of Virus vulnerability: Virus normally attached to various type of documents or storage medium, and is capable of destroying computer systems and networks.
- Bad mouthing attack: A third party malicious node establish a trust relationship and get benefits from it.
- Man-in-the-Middle attack: Malicious peers make independent connection with the victims and play messages between them, which make them believe that they are talking directly to each other over a private connection.
- Denial of Service attack: Each node should respond to every relevant query is likely to be an excessive burden on the peers which cause slowdown of the system.

D. Trust Model Used in P2P E-commerce

Trust plays a very important role in P2P networks for enabling peers to share resources and services credibly, and trust reflects a comprehensive evaluation of one user to another user's behaviours as well as ability. There are many ways to category trust models.

- According to different trust mechanism, the trust model can be classified into identity based trust model, role-based trust model, automated trust negotiation model, and reputation-based trust model.
  - Identity-based trust model: An operation cannot be authorized before authenticating the claiming entity. ex: PKI trust model.
  - Automated trust negotiation model: The architecture solves the problem of establishing trust sufficient to allow disclosure of sensitive credentials that must be exchanged to authorize application-level operations.
  - Reputation-based trust model: Reputation-based trust model allows establishing mutual trust in peers and reputation is a measure that is derived from direct or indirect knowledge on earlier interactions of agents.
- According to collection methods of trust value.
- According to algorithms that compute trust.

VIII. Conclusion

This paper has discussed some security related issues concerning authentication, authorization, confidentiality and non-repudiation in e-commerce. E-commerce security is the major issue keeping many commerce organizations afraid from using Internet for their business. Secured Socket Layer (SSL) and Secured Electronic Transactions (SET) are the major popular E-commerce security protocols. Each one of them has its domain of use, its products, its strategy, and its own encryption procedure. The traditional authentication mechanism is based on identity to provide security or access control methods; in addition, traditional encryption and authentication algorithm require high computing power of computer equipment. Therefore, how to improve the authentication mechanism and optimize the traditional encryption and authentication algorithm may be the focus of e-commerce. Effective trust models can facilitate in improving user trust in P2P e-commerce. Security related issues should be researched extensively for e-commerce. E-commerce security is the protection of e-commerce assets from unauthorized access, use, alteration, or destruction. Dimensions of e-commerce security; Integrity: prevention against unauthorized data modification, Confidentiality: protection against unauthorized data disclosure. Privacy: provision of data control and disclosure. Availability: prevention against data delays or removal. In conclusion, we believe that as long as the security issues are adequately addressed, the e-commerce should be able to achieve great success in the near future.

References

[1] Study of Security Problems in P2P E-commerce - LIN Ying School of Software Yunnan University Kunming,Yunnan Province, China,linying_yn@126.com

[2] E-COMMERCE SECURITY THREATS AND SOLUTIONS - Kalpana Sudam Dhende M. E. Society's Institute of Management and Career Courses (IMCC) 131, Mayur Colony, Kothrud, PUNE 411 029 (Maharashtra State) INDIA ksd.imcc@mespune.in
[3] The Study of E-Commerce Security Issues and Solutions - Dhurub Saxena1 ,Priyank Singhal2 1MCA Student, College of Computing Sciences and Information Technology, TMU, Moradabad 2AssistantProfessor, College of Computing Sciences and Information Technology, TMU, Moradabad
[4] Analysis of E-Commerce Security Protocols SSL and SET - Neetu Kawatra, Vijay Kumar Dept. of Computer Science Guru Nanak Khalsa College Karnal – India
[5] Enhancement of E-commerce System Security: A Review - VRINDA, MAMTA KATHURIA, AMITA ARORA YMCA University of Sciences and Technology, Faridabad
[6] E-Commerce- Study of Privacy, Trust and Security from Consumer's Perspective - Ms. Palak Gupta1 Dr. Akshat Dubey2 ¹Jagannath International Management School, New Delhi, India
[7] Rüdiger Schollmeier,"A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications", Proceedings of the First International Conference on Peer-to-Peer Computing, IEEE (2002).
[8] Electronic Payment System- ISA 767 (2008) Secure Electronic Commerce George Mason University
[9] "SET Secure Electronic Transaction LLC." Purchase, NY: SET Secure Electronic Transaction LLC, 2001. Available From Www.Setco.Org
[10] Marcus Goncalves. Industrial Networks Are Not Ready For E-Commerce. ARC Insights, Issue: 99-023, March 1999.
[11] D. K. Gangeshwer, " E-Commerce or Internet Marketing: A Business Review from Indian Context", International Journal of u- and e- Service, Science and Technology Vol.6, No.6 (2013), pp.187-194.
[12] Revathi C, Shanthi K, Saranya A.R "A Study on E-Commerce Security Issues" International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 12, December 2015.
[13] Niranjanamurthy, DR. Dharmendra Chahar "The study of E-Commerce Security Issues and Solutions" International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 7, July 2013.
[14] Forrester, US e-business Overview: 2003–2008, July 25, 2003.
[15] Online Advertising To Reach $33 Billion Worldwide By 2004. Forrester Research Press (1999) http://www.forrester.com/ER/Press/Release/0,1769,159,FF.html.
[16] Shazia Yasin, Khalid Haseeb. "Cryptography Based E-Commerce Security: A Review". IJCSI-Vol. 9, Issue 2, No 1, March 2012
[17] Randy C. Marchany, Joseph G. Tront, "E-Commerce Security Issues"Proceedings of the 35th Hawaii International Conference on System Sciences - 2002
[18] Anup K. Ghosh. Certifying E-Commerce Software For Security. National Institute For Standards And Technology (NIST), 1999.
[19] F.-Y. Leu, C.-H. Lin and A. Castiglione, —Special issue on cloud, wireless and e-commerce security‖, Journal of Ambient Intelligence and Humanized Computing, vol. 4, no. 2, (2013).
[20] Nikos Drakos. Security & Electronic Commerce:SSL Protocol. Security & Electronic Commerce Appendix, University of Leeds. 1997.

A Study on Architecture, Security and Applications of Cloud Computing

Ann Mary M B, Dilna K Daison and Mohidh K M
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Sajay K R
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: sajay.k.r@vidyaacademy.ac.in)

***Abstract*—The term "cloud computing" is a recent technology in the IT world. Cloud computing has formed the conceptual and infrastructural basis for tomorrow's computing. Cloud based services and service providers are being evolved which has resulted in a new business trend based on cloud technology. The global computing infrastructure is rapidly moving towards cloud based architecture. The main focus is on security issues in cloud computing that are associated with big data. Big data applications are a great benefit to the organizations, business, companies and many large scale and small scale industries. Cloud computing plays a very vital role in protecting data, applications and the related infrastructure with the help of policies, technologies, controls, and big data tools. Behind Though the term "Cloud Computing" is recent but the idea of centralizing computation and storage in distributed data centers maintained by third party companies. Cloud computing is aimed at providing IT as a service to the cloud users on-demand basis with greater flexibility, availability, reliability and scalability with utility computing model. This new paradigm of computing has an immense potential in it to be used in the field of industries, Banking sector, e-governance and in rural development.**

***Index Terms*—Cloud computing, cloud service, cloud security, computer network, distributed computing, machine learning, Banking, Business model,mobile computing.**

I. Introduction

Recent developments in the field of cloud computing have immensely changed the way of computing as well as the concept of computing resources. In a cloud based computing infrastructure, the resources are normally in someone else's premise or network and accessed remotely by the cloud users. In brief cloud is essentially a bunch of commodity computers networked together in same or different geographical locations, operating together to serve a number of customers with different need and workload on demand basis with the help of virtualization. Cloud services are provided to the cloud users as utility services like water, electricity, telephone using pay-as-you-use business model. These utility services are generally described as XaaS (X as a Service) where X can be Software or Platform or Infrastructure etc. Cloud users use these services provided by the cloud providers and build their applications in the internet and thus deliver them to their end users. So the cloud users don't have to worry about installing, maintaining hardware and software needed. And they also can afford these services as they have to pay as much they use. So the cloud users can reduce their expenditure and effort in the field of IT using cloud services instead of establishing IT infrastructure themselves.

Cloud computing has formed the conceptual and infrastructural basis for tomorrow's computing. The global computing infrastructure is rapidly moving towards cloud based architecture. While it is important to take advantages of cloud based computing by means of deploying it in diversified sectors, the security aspects in a cloud based computing environment remains at the core of interest. Cloud based services and service providers are being evolved which has resulted in a new business trend based on cloud technology. With the introduction of numerous cloud based services and geographically dispersed cloud service providers, sensitive information of different entities are normally stored in remote servers and locations with the possibilities of being exposed to unwanted parties in situations where the cloud servers storing those information are compromised. If security is not robust and consistent, the flexibility and advantages that cloud computing has to offer will have little credibility. This paper presents a review on the cloud computing concepts,type of cloud computing and different applications in cloud computing as well as security issues inherent within the context of cloud computing and cloud infrastructure.

A. Cloud Computing Basics

Cloud computing is a paradigm of distributed computing to provide the customers on-demand, utility based computing services. Cloud users can provide more reliable, available and updated services to their clients in turn. Cloud itself consists of physical machines in the data centers of cloud providers. Cloud Computing is the type of technology that provides remote services on the internet to manage, access, and store data rather than storing it on Servers or local drives. This technology is also known as Serverless technology. Here the data can be anything like Image, Audio, video, documents, files, etc.

B. Types of Cloud

There are three different types of cloud, private cloud, public cloud and hybrid cloud.

1) Private Cloud: This type of cloud is maintained within an organization and used for their internal purpose. So the utility model is not a big term in this scenario. Security, network bandwidth are not critical issues for private cloud.

2) Public Cloud: In this type an organization rents cloud services from cloud providers on-demand basis. Services provided to the users using utility computing model.

3) Hybrid Cloud: This type of cloud is composed of multiple internal or external cloud. This is the scenario when an organization moves to public cloud computing domain from its internal private cloud.

C. Cloud Computing Infrastructure

Cloud computing infrastructure is the collection of hardware and software elements needed to enable cloud computing. It includes computing power, networking, and storage, as well as an interface for users to access their virtualized resources.

Cloud infrastructure is made up of at least 7 basic components, including user accounts, servers, storage systems, and networks. Cloud environments are dynamic, with short-lived resources created and terminated many times per day. This means each of these building blocks must be secured in an automated and systematic manner. Read on to learn best practices that can help you secure each of these components.

Cloud infrastructure security is the practice of securing resources deployed in a cloud environment and supporting systems. There are different services provided by the cloud like, Software as a Service, Platform as a Service or Infrastructure as a Service(IaaS), and Software as a Service (SaaS),Education as a Service(Eaas),Logistics as a Service(LaaS),Manufacturing as a Service(Maas). These services are available over the Internet in the whole world where the cloud acts as the single point of access for serving all customers.

Cloud service hierarchichy is based on which a cloud is perceived in the form of iaas, paas and saas from any cloud end-user's viewpoint. Microsoft Azure, IBM, Amazon aws, Oracle,

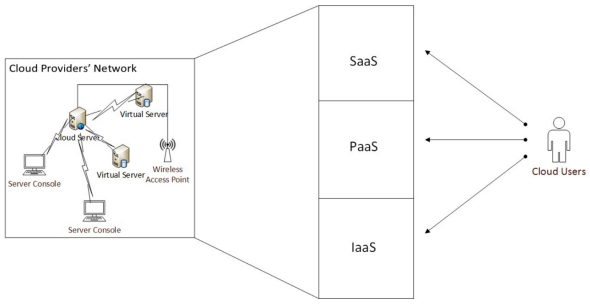


Fig. 1. cloud service hierarchy

Alibaba, and SAS are common cloud service providers. Most providers offer cloud infrastructure, communications services, data centers and data analytics.

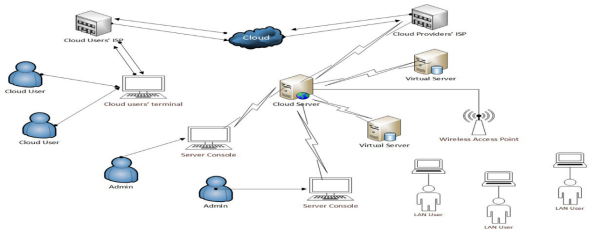


Fig. 2. Typical cloud architecture

II. Types of Cloud Computing

A. Mobile Cloud Computing

Mobile devices are becoming integral part of human life. The mobility feature of mobile devices changed the way that people use different technologies all over the world. The applications require high computing capabilities, big space to be stored, and maybe more security protection which are challenges for mobile devices. Cloud computing is the on-demand availability of computer system resources, especially data storage and computing power, without direct active management by the user. Cloud computing services provide wide use of clustering and virtualization of computing and memory resources of processing centers. Mobile Cloud Computing which is defined as a combination of mobile computing, cloud computing, and wireless network that come up together purpose such as rich computational resources to mobile users, network operators, as well as to cloud computing providers.Mobile Cloud Computing is meant to make it possible for rich mobile applications to be executed on a different number of mobile devices.

In this technology, data processing, and data storage happen outside of mobile devices. Mobile Cloud Computing has lots of advantages. Extended battery life, Improvement in data storage capacity and processing power, accessing data from anywhere, improved reliability and scalability are some of them. There are some problems associated with this. Protect the privacy of the users' data from unauthorized access and from malicious attacks is one of them. The integrity is a concern in which the data should not be altered or modified by intruders. Many cryptographic techniques can be used to provide solution to these information security concerns.

1) Architecture of Mobile Cloud Computing Technologies: The rapid increase of number of users on computing clouds loads the network which causes significant delays in delivery of processed information to the user. Cloud computing resources have to be located closer to the users in order to eliminate this deficiency. Computing cloud service users use two modes: offline and online. In offline mode the issues to the cloud and obtains the result after a certain period of time. In this mode, there is no direct connection between the user's computer and the cloud.In online mode user and cloud are always connected. And it will terminate only after the problem solution is complete.Mobile computing cloud technologies are used to eliminate the network load. Mobile

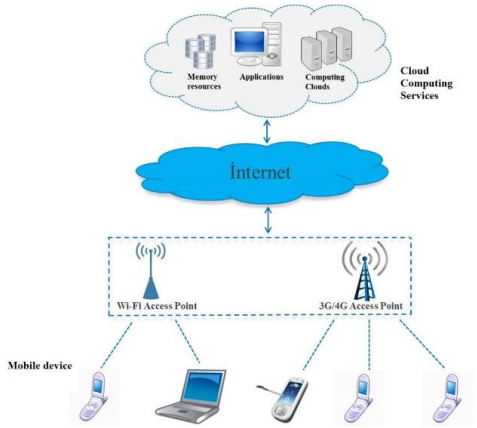


Fig. 3. Architecture of mobile computing cloud

computing cloud architecture consists of several components: mobile users, mobile connection operators, wireless connection devices, internet provider (ISP), and computing cloud providers. Mobile users connect to Cloud Computing over internet using base stations (Mobile connection, Wi-Fi Access Point). Such connections cause high load of the networks and delays in provision of processed information. These services are expensive. In the hierarchically structured architecture,

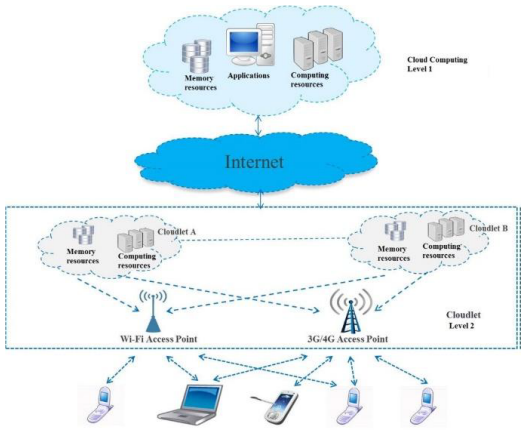


Fig. 4. Architecture of hierarchically structured mobile computing cloud

Mobile internet users load the necessary applications to the nearest cloudlet system server. This frees the internet network from overloading. Cloudlets (small computing clouds) are devices (server) located close to the users and provide faster delivery of necessary information to the users from the central servers. Cloudlets are located on the base stations on the 2nd level. Information can be re-loaded to the central servers.

2) Advantages Of Mobile Cloud Computing:

- Increasing battery life of mobile devices.
- Increasing computing and memory resources of mobile devices.
- Storage of user data on several back-up computers increases reliability.
- Reduces information loss risk on mobile devices.
- Increased scaling.

3) Problems in Mobile Cloud Computing:

- Mobile device limitations: Despite processing and memory resources of mobile devices there are important problems in use of energy sources and exploitation of complex applications.
- Connection quality: unlike wired networks using physical connection tools providing stable coverage area, the information transmission environment in mobile cloud computing environment continuously changes.
- Division of applications: due to limited resource issues, mobile devices cannot use software that does not require large computing and memory resources or consumption of large amounts of energy is required during use of these types of software.

4) Application Fields Mobile Computing Cloud Technologies:

- Mobile trade: Trade operations can be implemented using mobile devices. Mobile users can carry out mobile banking, mobile broker operations, mobile marketing and advertising operations via mobile trade.
- Mobile training: Conduction of trainings through mobile devices. This structure is based on e-training and mobility. Traditional mobile training programs have limitations due to high network services, weak network speed and limited training resources.
- Mobile games: As processing capabilities of the mobile devices are limited, they use very basic games. More complex, appealing games use cloud technologies. This provides advantages such as reducing energy use, faster use of game software due to processing power of the clouds.
- Mobile healthcare: Medical applications existing in the mobile environment are called mobile medical applications and used for medical treatment, patient follow-up, operative diagnosis, observation of the patient at any place or time via internet, alarm system for measurement of pulse, blood pressure and alcohol level in emergency situations etc

5) Mobile Cloud Computing in Healthcare:

Healthcare applications require large amounts of computational and communication resources, and involve dynamic access to large amounts of data within and outside the health organization leading to the need for networked healthcare. Mobile cloud computing could provide the necessary computational resources at the right place and right time through cloudlet and fog computing-based architectures. Moreover, big data and relevant technologies could provide the data management and analytics solutions that are necessary to reduce healthcare costs and improve system and clinical inefficiencies. The massive growth of information and communications technology (ICT) is also visible in health informatics, bioengi-

neering and Healthcare Information Systems (HIS). With the help of analytical tools and decision support tools healthcare sector can have a big change. This will help to reduce the health care costs, system inefficiencies, variations in quality of care, high prevalence of medical errors, greater public analysis of government spending. There were different studies in healthcare applications. A quantitative modeling study to demonstrate the potential of computational grids for its use in healthcare organizations to deploy diverse medical applications was presented. The future networks connecting healthcare systems are likely to be dominated by the analytics applications that require frequent, low-latency, communications. A comparative study on mobile computing to get a better solution for mobile healthcare applications was also presented. A mobile cloud architecture relevant to healthcare applications that stores and manages personal healthcare data was carried out. Many studies have explored the networked systems and QoS in transferring data over different networks, which is very important in many applications especially in healthcare. Networked healthcare systems will eventually be designed as sustainable enterprise systems which will be part of networked smart city information and operations systems.

Infrastructure for Health Care Big Data:

The traditional infrastructures involve set of cloud resources accessed remotely by the users of different types of devices via through the Internet. The massive spread of mobile applications in all and every area of the people's life resulted in huge amounts of data that need to be processed and analyzed efficiently in less time and power complexity which imposes the need for new competitive MCC models other than the traditional one. Performance Enhancement Framework using the Cloudlet can be considered as a closer cloud with many advantages and capabilities to avoid several limitations of distant cloud. The researchers introduced an efficient cloudlet MCC model in which the mobile users communicate directly to the cloudlet instead of the enterprise cloud. Their model can be applied in many environments including hospitals were big amounts of data need to be saved and processed.

The Big Data is a recent term associated with the huge amounts of stored obtained data due to the revolutionary advances in different technologies including: cloud computing, spread of social media, and wireless communication technologies. It is defined according to: the size of data (volume), types of data based on the producing source (variety), and the time frequency to generate the data (velocity); every, minute, day, month, or a decade. Some applications needs real time processing for this data such as health applications where the data analysis and extracting the right decisions makes a difference between patients life and death. In this MCC model the cloudlets are placed nearby the hospital and cover an area that can be accessed by authorized people who can access the patients' information and follow their status remotely.

6) Data Analysis:

The science of examining raw data with the purpose of drawing conclusion or inspecting, leaning, modelling and transforming data with the purpose of highlighting useful information is called data analytics. Data analytics focuses on the conclusion reached on the basis of evidence and reasoning, the method of deriving a result based solely on what is already known by the researcher. Data analytics science is divided into exploratory data analysis (EDA), conrmatory data analysis (CDA) and qualitative data analysis (QDA). In EDA new features in the data are discovered, in CDA existing hypotheses are proven true or false and QDA is being used in social sciences to draw conclusion from non-numerical data for example, photographs, videos or words. Developers developed a Cloud data analytics service which is based on Daytona. Daytona is an interactive MapReduce runtime optimized for data analytics. In their model, Excel and other client application provide the data entry and other interaction interface to the user, and bridges the gap between the client and Cloud.

7) Tools for Data Analysis:

There are different tools available for data visualization and analysis,

- Data Wrangler: It is designed for cleaning and rearranging data.
- Google Refine: it can import and export data in different formats such as XML, Excel and JSON files.
- The R Project: It is a general statistical analysis platform which runs on command line.
- Google Fusion Tables: It is one of the simplest ways to turn data into chart or map.
- Tableau Public: It can turn data into any number of visualizations, like from simple to complex.
- VIDI: It is a tool for Drupal content management system.
- Zoho Reports: This tool can take data from number of file formats or directly from a database and then turn it into tables, charts and pivot tables and other formats which are familiar to spread sheet users.
- Choosel: A open-source web-based framework which is designed for charts, graphs, clouds, maps and timelines
- Google Chart Tools: It is a full edge self-contained application for uploading and sorting data, and also generating maps and charts
- Quantum GIS (QGIS): This tool is designed for creating maps that offer sophisticated, detailed data base analysis of geographic regions
- Gephi: It is photo shop for data and this open-source beta project is designed for visualizing statistical information

Talend Open Studio provide Hadoop technologies for big data transfer and big data analytics. Map/reduce a computational paradigm is implemented by Hadoop, where application is divided into various pieces of work and each of which may be executed or- re executed on any node in the cluster. Weka is another tool, which is a collection of machine learning algorithms for data mining. Weka provide tools for data preprocessing, regression, classification, clustering, visualization and association rules. Rapid Analytics is open source and is one of most widely used data mining and predictive analysis solutions worldwide, it is built around RapidMiner.

Here, we discussed networked healthcare systems and the role that mobile cloud computing and big data analytics play in its enablement. A Cloudlet based Mobile Cloud Computing infrastructure to be used for healthcare big data applications was described as it require large amounts of computational and communication resources, and involve dynamic access to large amounts of data within and outside the health organization and this can be implemented efficiently via cloud and mobile cloud systems.

B. Green Cloud Computing

Green cloud computing is a coined term that means making the practices and approaches of using technological advances like computing and other IT resources sustainable for potential environmental benefits.

The growing number of companies around the world makes a significant impact on the environment. This growth means for a great need to use data centers, an increase in the daily commute of workers, and an influx of office materials and supplies needed on a daily bases. Green cloud computing answers these looming environment issues by providing options that can lower emitted carbon footprints around the world.

Many companies are now leaning on the cloud services as such many technological applications and practices that can reduce environmental impacts are being developed daily. Green cloud computing makes it possible to maintain and enhance business operations and processes while looking after the environment.

III. Applications of Cloud Computing

A. Transformation in Development of Industries using Cloud Computing

In industries, the high costs of infrastructure maintenance and equipment procurement may influence business revenue. So by using cloud computing services, a user can access virtual resources instead of buying physical infrastructure and computing resources. Such services provide a configurable and extensible infrastructure that supports different types of inputs (e.g., customer needs. device control, and sensors) and outputs (e.g., indicators, analyzer, and actuators).

Cloud computing services, such as IaaS, PaaS and SaaS, provide cloud-based applications to users who perform tasks using the cloud. Microsoft, IBM, Amazon, Oracle, Alibaba, and SAS are common cloud service providers. Most providers offer cloud infrastructure, communications services, data centers and data analytics. Data analytics is essential for cloud computing. Microsoft, Gartner etc classify data analytics into four as follows: descriptive, diagnostic, predictive and prescriptive.

The cloud computing has developed new services for different industries:

- Everything As A Service (XAAS): It is a cloud computing service that involves various services through a network. It is concerned with social science, which aims to enhance co-creation and value creation. And also used for data access and communication. It has been applied to marketing, e-learning, manufacturing.
- Education As A Service (EAAS): It is a business model of a delivery system that unbundles the components of higher education and offers students the option to pay for only those courses that they want and need. It remains extensible, it can be scaled to fit user's requirement and has been applied in medical and power education.
- Logistics As A Service (LAAS): It is prevalent in the manufacturing, transportation and international trade sectors. It is a cloud computing service that involves various services through a network.
- Manufacturing As A Service (MAAS): It combines advanced technologies and service oriented technologies. It is a cloud computing service that involves various services through a network. It is aiding the evolution of manufacturing processes and equipment maintenance.

In this study investigated EaaS, LaaS and Maas by applying a ML model to predict UCI data from various sources.ML algorithms were used to solve and integrate these issues into different cloud computing solutions to achieve the goals of EaaS, LaaS, and MaaS.

We use these methods to analyse:

1) Procedure of Knowledge Discovery in Databases (KDD):

It is the process of acquiring potential useful information by analyzing or mining data collections. It is widely used technique of data analytics is based on procedures that analyze data from different sources. The standard procedures are selecting a target dataset, pre-processing data, splitting the data ,training model and evaluation. To achieve the goal of cloud computing ,this study is done by Microsoft Azure ML Studio.

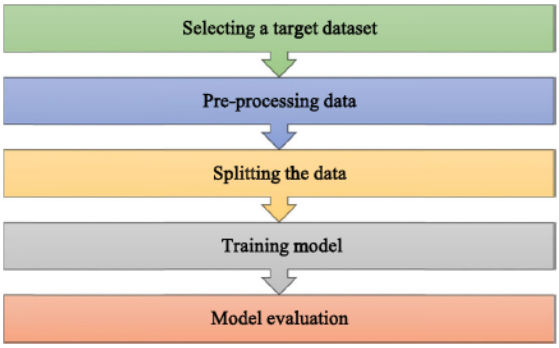


Fig. 5. The procedure of KDD

2) Machine Learning Modeling:

This study supports SVM, tree structure algorithm, and classifier-related algorithm were employed in predictive analytics. It also used ensemble learning to enhance the accuracy of the algorithm.

- Two-Class Boosted Decision Tree: It is an example of ensemble learning and relies on the ensemble of trees to forecast. Succeeding trees adjust the errors of preceding

trees. The algorithm is repeated until it reaches a terminal condition.

- Two-Class SVM: It is based on an SVM algorithm used to analyze input data and recognize patterns in a multi-dimensional feature space.
- Two-Class Logistics Regression: It is a supervisory, learning algorithm that predicts the probability of a binary event's occurrence using a logistic function.
- Two-Class Bayes Point Machine: It approximates the Bayes-optimal decision of linear classifiers by choosing a generalization performance classifier.It approximates in center of solution space so not prone to overfitting the training data.
- Two-Class Decision Forest: It is an ensemble learning algorithm intended for classification tasks. It creating multiple related models and combining them in some way to enhance performance.

3) Procedure for Data Analytics :

In this study, the procedure for data analytics involved importing open data, preprocessing and selection , training the data, evaluating the model ,and formulating accurate solutions from observed results. We take Data sources from UCI ML

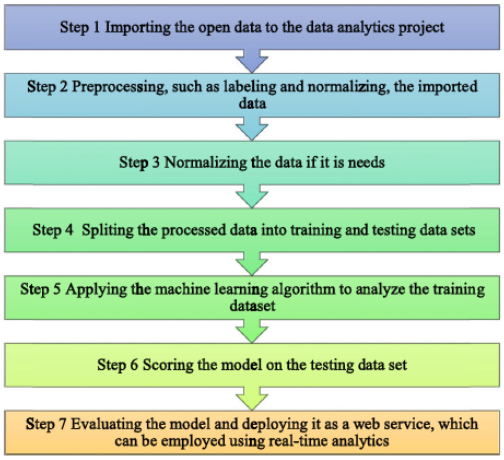

Step 1 Importing the open data to the data analytics project
Step 2 Preprocessing, such as labeling and normalizing, the imported data
Step 3 Normalizing the data if it is needs
Step 4 Spliting the processed data into training and testing data sets
Step 5 Applying the machine learning algorithm to analyze the training dataset
Step 6 Scoring the model on the testing data set
Step 7 Evaluating the model and deploying it as a web service, which can be employed using real-time analytics

Fig. 6. The procedure of the data analytics

repository. Table 1 exhibits the dataset description. And we get the results by this study applied cloud computing to education, logistics and manufacturing. The result shows how Microsoft Azure ML studio could assist developers in creating a cloud ML model for these industries.

We have Evaluation Criteria , this study used accuracy, the receiver operating characterstic curve (ROC),and the area under the curve of the ROC(AUC) to evaluate the analytics.

- Prediction of student's learning performance using EaaS: It is contains application of educational data mining to predict students learning performance. We use normalization and by using this method teacher can predict the students 's performance and create an adaptive learning strategy on the cloud.

Dataset	Description	Instances
Educational dataset	The data comprising thirty-three attributes, e.g., grades, social related features, and self perception of learning.	649
Logistical dataset	The data has seven attributes, including the size of luggage boot, buying price,number of doors, passenger capacities, safety of the car, car acceptability.	1728
Manufacturing dataset	The data has 591 features containing method, classifications, and time stamp for each instance.	1567

TABLE I
THE LIST OF DATASET

- Vehicle Selection using LaaS: The transportation tool plays an important role in logistics and some studies were dedicated to solving car rental logistics. These classification algorithms were employed to predict vehicle acceptability
- In Yield Failure Prediction in Semiconductor Manufacturing Process using MaaS: The manufacturing case aimed to forecast yield failure predictions for a semiconductor manufacturing process. And it can assist users in predicting the remaining life of equipment, reduce unplanned forced outages, and minimize the risk of catastrophic failures

B. Banking Sector using Cloud Computing

Cloud computing today encompasses every vertical in the market across sectors. Organizations are adopting innovative cloud apps to support their everyday business operations.
The dramatic changes taking place in banking require new ways to maximize profitability and returns. Cloud technology offers secure deployment options that can help banks develop new customer experiences, enable effective collaboration and improve speed to market—all while increasing IT efficiency. Banks that take advantage of cloud computing are better positioned to respond to economic uncertainties, interconnected global financial systems and demanding customers.
They can use information to enhance customer segmentation techniques. Perceived cost savings, ease of scaling-in and scaling-out, faster time to-market for deploying systems, virtualization of enterprise-wide data as a service, enterprise technology standardization, and the ability to access data and applications on the move are all critical consideration factors that can drive financial services firms to adopt cloud computing.
The context-driven Variability provided using cloud computing makes it possible for banks to personalize customer interactions and adapt to subtle changes, which leads to a more user-centric experience. With cloud computing, various financial institutions only have to budget for functional expenses and wage for the services they use. This makes it effortless and

more cost effective to test new applications on the cloud versus prevailing conventional infrastructures. No cloud computing services model is customary to meet all the technology requirements for every financial organization. Banks should develop and preserve an application portfolio consisting of both cloud and on- ground applications.

Cloud service models offer financial organization the option to minimize operational wage. The key to achievement lies in choosing the right cloud services model to meet business needs.

1) Cloud Operating Models:

The third aspect of choosing the right cloud services delivery model is determining the appropriate operating model for the required mix of resources and assets.

- Staff augmentation: Financial firms can gain cloud expertise by hiring people with the right skill sets from service vendors. This operating model allows for flexibility and lets firms choose the best resource for each specific requirement.
- Virtual captives: Virtual captives have a dedicated pool of resources or centers to help with cloud operations and meet demand. This operating model is a good alternative to a complete outsourcing approach.
- Outsourcing vendors: This approach uses offshore centers, facilities, and people from a third party vendor to handle cloud operations. The model combines resources and investments to cater to cloud services for multiple banks.

2) Advantages:

- Cost Savings: Business sharpness is determined by the cost an organization incurs. Billing is a non-core process for banks, and outsourcing it to a less expensive mediator allows them to route their capital into core technology-based functions.
- Scalability: If well designed, cloud solutions empower banks to meet customer demands and scale quickly, dynamic provisioning of computing resources, will save business users and IT experts from engineering the systems for peak loads.
- Time to market: With cloud computing, time to market can be curtailed from months to weeks or days, depending on the size of a bank.A self-service based, on-demand and real-time monitored cloud helps by: Accelerating computing power for when current applications need to deal with peak loads.
- Data Virtualization: It is the assimilation of data from multiple and diverse sources across the enterprise or external sources for the on-demand consumption by a wide range of applications in a virtualized manner. Accordingly, accessing that data from a single virtual source would drive scores of data consolidation within banks.
- Mobility: Likewise, they want similar interfaces for banking services-specific applications. And since a cloud facilitates users to access systems and infrastructure using a web browser or customized clients regardless of location and time, advancement of such interfaces has started taking shape.

3) Challenges:

When a bank moves into cloud computing, there are two prime challenges that must be addressed:

- Security: The confidentiality and security of commercial and personal data and mission-critical applications is pre-eminent. Banks cannot allow the danger of a security breach. Ultimately, for cloud computing to gain full acceptance within the banking services sector, cloud services must be harmlessly integrated into existing security platforms and processes.
- Regulatory and compliance: Many banking mangers require that financial data for banking consumers stay in their native country. Certain compliance arrangements require that data not to be mixed with other data such as on shared servers or databases. As a result banks must have a fair understanding of where their data is stored in the cloud.

Security issues which cloud clients should advert are:

- Privileged user access: There dwell sensitive data that is processed outside the organization inherent risk of security of data because outsourced services bypass the physical and logical IT controls.
- Regulatory compliance: Customers are responsible for the security of their data. Traditional service providers are subjected to external audits and security certifications.
- Data location: When users use the cloud, they have no knowledge about the hosted data. Distributed data storage is a main reason of cloud providers that can cause lack of control and that is risky for customers.
- Data segregation: As cloud is typically in a shared environment in that data can be shared. So there is the danger for data loss. Is encryption available at all phases, and were these encryption patterns designed and tested by experienced professionals.
- Recovery: It is very essential to recover the data when some problem occurs and creates failure. So, the main question arises here is that can cloud provider restore data completely or not, this issue can cause a stalemate in security.
- Investigative support: Cloud technology services are difficult to investigate, because logging and data for multiple customers may be co-located and may also be spread across an ever-changing set of hosts and data centers.
- Long-term viability: Cloud computing provider will never go broke or get acquired by a larger company with maybe new policies.

IV. Security Issues in Cloud Computing

- DDOS (distributed denial of service) attack is one common yet major attack for cloud computing infrastructure.
- Data segregation and session hijacking are two potential and unavoidable security threats for cloud users.

- Data loss: It is one of the issues faced in cloud computing. security of cloud service is to break by hackers then it may be possible that hackers will get access to our sensitive data or personal files.
- User account hijacking: Hijacking is the most serious security issue in cloud computing. if somehow the account of user or an organization is hijacked by hacker.
- Denial of service (dos) attack: This type of attack occurs when the system receives too much traffic. mostly dos attacks occur in large organizations such as the banking sector, government sector, etc. when a dos attack occurs data is lost.

A. Authentication in Cloud

Cloud based authentication allows authorized users across networks and continents to securely access information stored in the cloud with authentication provided Through cloud-based services. cloud-based authentication is to protect companies from hackers trying to steal confidential information. Cloud computing is changing the way we interact with devices, software, data and processes. Identity management and authentication form the basis for security whether in the cloud or on the local network.

Security is the most prioritized aspect for any form of computing. cloud computing approach could be associated with having users' sensitive data stored both at clients' end as well as in cloud servers, identity management and authentication are very crucial in cloud computing.

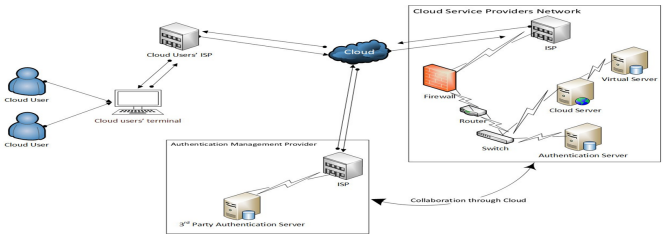


Fig. 7. Authentication in the Cloud

B. Security with Big Data in Cloud Computing

In recent years, big data and cloud computing are the major issues in an organizations. it enables computing resources to be provided as information technology service with high efficiency and effectiveness. today's big data is the one of the major problem that researchers try to solve it. the main focus is on security issues in cloud computing that are associated with big data. big data applications are a great benefit to organizations, business, companies and many large scale and small scale industries.

Big data is the term for data sets, so large and complicated that it becomes difficult to process using traditional data management tools or processing applications. Data and to identify patterns it is very important to securely store, manage and share large amounts of complex data. Big data is the word used to describe massive volumes of structured and unstructured data that are so large that it is very difficult to process this data using traditional databases and software technologies. In many organizations, the deployment of big data for fraud detection is very attractive and useful. The challenge of detecting and preventing advanced threats and malicious intruders, must be solved using big data style analysis.

These techniques help in detecting the threats in the early stages using more sophisticated pattern analysis and analyzing multiple data sources. Not only security but also data privacy challenges existing industries and federal organizations. With the increase in the use of big data in business, many companies are wrestling with privacy issues. Data privacy is a liability, thus companies must be on privacy defensive. But unlike security, privacy should be considered as an asset, therefore it becomes a selling point for both customers and other stakeholders. There should be a balance between data privacy and national security. Big data is one of the most talked about technology. security and privacy issues are magnified by the three v's of big data: velocity, volume, and variety.

V. Conclusion

Cloud computing has a growing market that many enterprises and organizations are adopting to enable their digital transformation and we also discussed not only cloud computing Technologies but also current cloud computing services research such as EaaS, LaaS, and MaaS and also including security issues, applications, cloud computing types etc. And cloud computing within the banking sector will require vendors and banks to overcome its listed challenges together. Private clouds are expected to increasingly become the deployment model for cloud services among banks, giving financial institutions full control through ownership and operations of their cloud systems.

And we reviewed mobile computing cloud architecture, their application areas and advantages of hierarchically structured system and also problems occurring during the use of mobile computing clouds are studied and solution methods are indicated. Through Mobile cloud computing, it maximizes the utilization of mobile devices capabilities to run intensive-computing applications. These intensive jobs are executed in the mobile cloud computing infrastructure overcoming the mobile device limitations, saving energy, and providing better throughput. A Cloudlet–based Mobile Cloud Computing infrastructure to be used for healthcare big data applications was described. The techniques, tools, and applications of big data analytics were reviewed.

The security issues in cloud computing are somewhat sensitive and crucial on the basis of sociological and technological viewpoints – the technological inconsistency that results in security breach in cloud computing might lead to significant sociological impacts. Cloud environment is widely used in industry and research aspects; therefore, security is an important aspect for organizations running on these cloud environments. Using proposed approaches, cloud environments can be secured for complex business operations.

REFERENCES

[1] Monjur Ahmed, Mohammad Ashraf Hossain, "Cloud Computing And Security Issues In The Cloud," International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.1, January 2014.

[2] M Ramya, P Sonali, N Naga Dharshini, "Security With Big Data In Cloud Computing," Journal of Analysis and Computation (JAC),International Conference on Emerging Trends in IOT & Machine Learning, 2018.

[3] Yu-Hsin Hung, "Investigating How the Cloud Computing Transforms the Development of Industries," Special Section On Distributed Computing Infrastructure For Cyber-Physical Systems , Vol.7, pp. 181506 - 181517, December, 2019.

[4] K Sudhakar, G Vinay Kumar and L Sudha Rani, "A View on Cloud Computing in the Banking Sector ," International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 3305 - 3308, 2014.

[5] vi. Alakbarov R G, Alakbarov O R , "Mobile Clouds Computing: Current State, Architecture and Problems," Second International Conference on Electrical, Computer and Communication Technologies (ICECCT), February, 2017 .

[6] Lo'ai A. Tawalbeh, Rashid Mehmood, Elhadj Benkhlifa, Houbing Song, " Mobile Cloud Computing Model and Big Data Analysis for Healthcare Applications," Special Section On Healthcare Big Data, October, 2016.

[7] Q Zhang, L Cheng, and R Boutaba, " Cloud computing: State-of-the art and research challenges," J. Internet Services Appl., vol. 1, no. 1,pp. 718, 2010.

[8] M Armbrust, A Fox, R Grifth, A. D Joseph, R Katz, A Konwinski, G Lee, D Patterson, A Rabkin, I Stoica, M Zaharia, "A View Of Cloud Computing," Commun. ACM, vol. 53, no. 4, p. 50, 2010.

[9] D West, "How Mobile Devices Are Transforming Healthcare," Issues Technol. Innov., vol. 18, no. 1, pp. 1–11, 2012.

[10] M P Vaishnnave, K Suganya Devi, P Srinivasan, "A Survey on Cloud Computing and Hybrid cloud," International Journal of Applied Engineering Research ISSN 0973-4562 Vol.14, Number 2 , pp. 429-43, 2019.

[11] L Kumar , Nishant Malik , Gourav Agghi , Ajay Anand, "Mobile Cloud Computing," IJRIT International Journal of Research in Information Technology, Vol. 2, no. 9, pp. 787-792, 2014.

[12] L Xu, E Xu, and L Li, "Industry 4.0: State of the art and future trends," Int. J. Prod. Res., vol. 56, no. 8, pp. 2941-2962, 2018.

[13] Bisong, A Rahman, S.S.M., " An Overview of the Security Concerns in Enterprise Cloud Computing," International Journal of Network Security & Its Applications, Vol.3, no.1, pp. 30-45, 2011.

[14] Suresh K S and Prasad K V, "Security Issues and Security Algorithms in Cloud Computing," International Journal of Advanced Research in Computer Science and Software Engineering, Vol.2, no.10, pp. 110-114, 2012.

Wireless Sensor Network : Deployment and Security Issues

M Shahil Haris, Praseetha Praveen and Suraj Kumar M
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Manesh D
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: manesh.d@vidyaacademy.ac.in)

***Abstract*—A wireless sensor network (WSN) is a network of multiple small inexpensive smart sensor nodes that are capable of physical and environmental data acquisition and wireless communications. The key technology of the WSN network layer is the routing protocol. Among this field, the clustering routing protocol has become a research hotspot due to the high scalability and obvious advantages in saving node energy consumption to extend WSNs' lifetime. To solve the problem of 'hot spot', a routing algorithm named EEUC was proposed. Software defined networking is to ensure proper working order and network efficiency of a network of sensor nodes. It represents a new paradigm to simplify the management and configuration of networks by providing high-level abstractions for networking ideas. With the help of various localization techniques, an energy efficient Wireless Sensor Network is obtained. Threats such as Denial of service (DoS) attacks , Data theft and Distributed Denial of Service(DDoS) are some major security issues faced by Software Defined Networking . This paper proposes the method FRESCO to solve the above mentioned security threats.**

***Index Terms*—Motes, Hotspot, Multi-hop routing, EEUC, FM-CUR, Base station, FRESCO**

I. Introduction

WSN (Wireless Sensor Network) is the most standard services employed in commercial and industrial applications, because of its technical development in a processor, communication, and low-power usage of embedded computing devices. The wireless sensor network architecture is built with nodes that are used to observe the surroundings like temperature, humidity, pressure, position, vibration, sound, etc. These nodes can be used in various real-time applications to perform various tasks like smart detecting, a discovery of neighbor nodes, data processing and storage, data collection, target tracking, monitor and controlling, synchronization, node localization, and effective routing between the base station and nodes. Presently, WSNs are beginning to be organized in an enhanced step. It is not awkward to expect that in 10 to 15 years that the world will be protected with WSNs with entree to them via the Internet. This can be measured as the Internet becoming a physical network. This technology is thrilling with infinite potential for many application areas like medical, environmental, transportation, military, entertainment, homeland defense, crisis management, and also smart spaces.

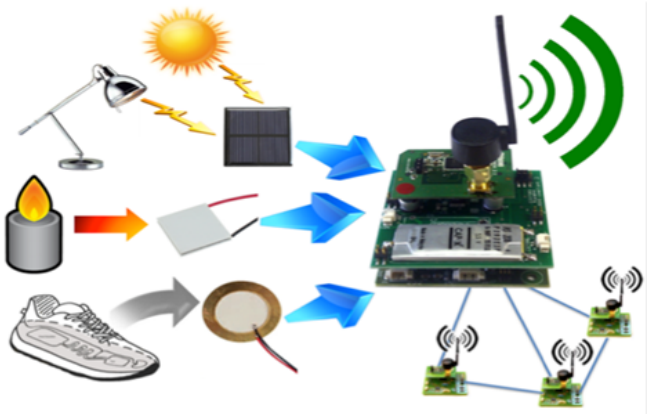


Fig. 1. Wireless Sensor Network

II. Wireless Sensor Network

A Wireless Sensor Network (WSN) is a network of multiple small inexpensive smart sensor nodes that are capable of physical and environmental data acquisition and wireless communications. A sensor node can be defined as a device which has the parts such as a power/battery unit, a communication module and a sensing unit. These nodes typically have very limited resources and should always be used with energy efficiency in mind. Nodes are tiny computers, which work jointly to form networks. The sensor node is a multi-functional, energy-efficient wireless device also called as motes. The applications of motes in industrial are widespread. A collection of sensor nodes collects the data from the surroundings to achieve specific application objectives. The communication between motes can be done with each other using transceivers. In a wireless sensor network, the number of motes can be in the order of hundreds/ even thousands.

A. Characteristics of Wireless Sensor Network

The characteristics of WSN include the following:

- Network arrangements can be carried out without immovable infrastructure.
- Apt for the non-reachable places like mountains, over the sea, rural areas, and deep forests.
- Flexible if there is a casual situation when an additional workstation is required.
- Execution pricing is inexpensive.
- It avoids plenty of wiring.
- It might provide accommodations for the new devices at any time.
- It can be opened by using centralized monitoring.

B. Advantages of Wireless Sensor Networks

The advantages of WSN include the following:

- The consumption of Power limits for nodes with batteries
- Capacity to handle node failures
- Some mobility of nodes and Heterogeneity of nodes
- Scalability to a large scale of distribution
- Capability to ensure strict environmental conditions
- Simple to use
- Cross-layer design

III. Wireless Sensor Network Applications

Wireless sensor networks may comprise numerous different types of sensors like low sampling rate, seismic, magnetic, thermal, visual, infrared, radar, and acoustic, which are clever to monitor a wide range of ambient situations. Sensor nodes are used for constant sensing, event ID, event detection and local control of actuators. The applications of wireless sensor networks mainly include health, military, environmental, home, and other commercial areas.

- Military Applications
- Health Applications
- Environmental Applications
- Home Applications
- Commercial Applications
- Area monitoring
- Health care monitoring
- Environmental/Earth sensings
- Air pollution monitoring
- Forest fire detection
- Landslide detection
- Water quality monitoring
- Industrial monitoring

IV. Wireless Sensor Network Architecture

The most common wireless sensor network architecture follows the OSI architecture Model. The architecture of the WSN includes five layers and three cross layers. Mostly in sensor network, we require five layers, namely application, transport, network, data link and physical layer. The three cross planes are namely power management, mobility management, and task management. These layers of the WSN are used to accomplish the network and make the sensors work together in order to raise the complete efficiency of the network.

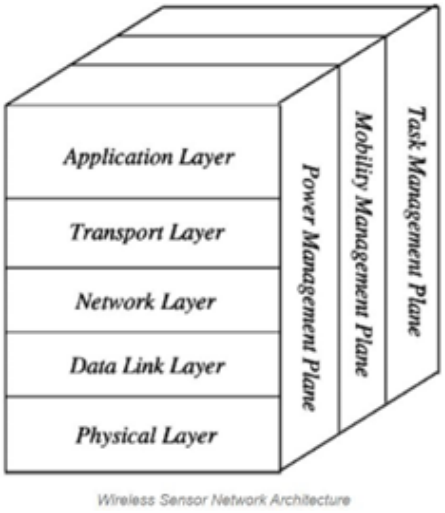


Fig. 2. Wireless Sensor Network Architecture

The key technology of the WSN network layer is the routing protocol. Among this field, the clustering routing protocol has become a research hotspot due to the high scalability and obvious advantages in saving node energy consumption to extend WSNs' lifetime. The single-hop data transmission between the cluster head nodes and the base station cannot meet the technological requirements of the long-distance communication.The use of multi-hop routing between cluster heads brings a new problem of unbalanced energy consumption. The cluster head node consumes relatively large energy due to its multi-tasks including manage and control of the intra-cluster member nodes, collect the data from member nodes, actualize data fusion and forward to the aggregation node. . In the clustering routing protocol that combines multi-hop routes between clusters, the cluster head node also forward the data from other cluster heads, which leads to the heavier forwarding tasks of the cluster heads closer to the sink node and the denser traffic distribution. This issue is called the "hot spot" of the network . Excessive consumption of some cluster head nodes reduces the energy efficiency of the entire network. To solve the problem of 'hot spot', a routing algorithm named EEUC was proposed.

V. Energy - Efficient Unequal Clustering Algorithm

The basic principle of energy - efficient unequal clustering algorithm is as follows: Firstly, determine the size of the competition radius of each node according to the distance from the node to the aggregation node, leading the nodes closer to the convergence point have a smaller competition radius. Then, member nodes with fewer burdens have more energy to use for forwarding data of other cluster head nodes which achieves the goal of balanced node energy consumption. Simulation results show that the unequal clustering routing protocol can

effectively solve the energy imbalance problem of multi-hop routing between clusters and prolong the life cycle of the network.
The unequal clustering routing EEUC algorithm is studied and improved :

1) In the process of cluster head election, the negotiation mechanism is replaced by time competition to reduce the burden of network traffic.
2) The distance between node and sink node is added as the cluster head elect parameter, which makes the cluster head node in "hot spot" are more dense and enhances the unequal distribution of clusters within the network area.
3) Ordinary nodes select cluster heads based on the competition radius and distance between nodes, which reasonably plan the cluster size and further balances the network energy consumption.
   The improved algorithm reduces network energy consumption and extends the network life cycle with higher stability. In the improved algorithm, the RSSI mechanism is employed to obtain the distance between nodes. The ranging method meets the requirements of low power and low cost.Another method is to use Fuzzy-based Multi-hop Unequal Cluster Routing (FMUCR) algorithm to solve the problem of hotspot.. However, when it comes to practical applications, the result is not ideal because the ranging error is large. It is the result of the unstable electromagnetic wave due to factors such as environment, geographical location, obstacles, etc.

VI. Software Defined Networking

To ensure proper working order and network efficiency of a network of sensor nodes, an effective WSN management system has to be integrated. However, the inherent challenges of WSNs such as sensor/actuator heterogeneity, application dependency and resource constraints have led to challenges in implementing effective traditional WSN management. This difficulty in management increases as the WSN becomes larger. Software Defined Networking (SDN) provides a promising solution in flexible management WSNs by allowing the separation of the control logic from the sensor nodes/actuators. The advantage with this SDN-based management in WSNs is that it enables centralized control of the entire WSN making it simpler to deploy network-wide management protocols and applications on demand. It represents a new paradigm to simplify the management and configuration of networks. This is done by providing high-level abstractions for networking ideas. The network is split into three independent concepts namely:- Data plane (forwarding phase), Control plane and the application plane.
Data plane: It is responsible for packet forwarding. It is not responsible for generating the rules by which it operates but simply acts in a manner which its rules describe. To enable communication with the network, a device only needs to support the data plane and communication with the control plane through a south bound interface.
Control plane: It represents the abstraction of routing and decision making processes from the network. This allows for centralized control and management across the entire network. It has a global view across the network which allows for protocol management and protocol deployment to be done in one place.
Application plane: It represents the abstraction of network services and applications from the network.
Commonly used protocols: There are two major protocols which can be used to facilitate south bound communication namely: ForCes and Openflow. The ForCes was developed by the Internet Engineering Task Force (IETF). The Openflow, which was developed by OpenNetwork foundation, has greater market use at the moment and is considered as the standard go to protocol for south bound interface communication.

SDWSN (Software Defined Wireless Sensor Networks) The SDN's were created in an address centric environment whereas WSN's were created for a data centric environment, at this stage obtaining address for sensor nodes is proving to be a non-trivial solution. Openflow as mentioned before is a flow-based protocol which provides flow rules for devices in the data plane. The Openflow uses TCP as a transport layer protocol. TCP is inherently address centric, which suffers in environment where loss and link failures are common. SDWNwise is another method of enabling SDWSN's and it was mainly developed for two purposes:- to minimize the information exchange between the controller and sensor nodes and also to allow programming of sensor nodes over the network.

VII. Localisation

Localization is a process to compute the locations of wireless devices in a Wireless Sensor Network composed of a large number of inexpensive nodes that are densely deployed in a region of interests to measure certain phenomenon. Positional information enables the use of geolocation routing protocols. Geolocation routing can be useful to improve the efficiency of a network. This can be done by routing packets in such a manner that the least amount of energy is consumed by the radio transmitters. Generally localisation is done in the following steps:- position/distance estimation, position computation, error correction. From the localisation perspective nodes can be classified as:
Anchor: A node which is self-locating, knows its own position.
Unlocalised / Unknown: A node which does not know its own position.
Localised: A node which at some point was an unlocalised node but has had its position estimated.

VIII. Classification of Localisation Techniques

The localization techniques can be categorized into two basic types one is centralized and second is decentralized schemes or distributed techniques. The decentralized scheme is divide into Free-range-based and Range-base.

1) Centralized localization technique: In this technique the base station (BS) collects and calculates all the data and measurements. After the calculation of the results a reply

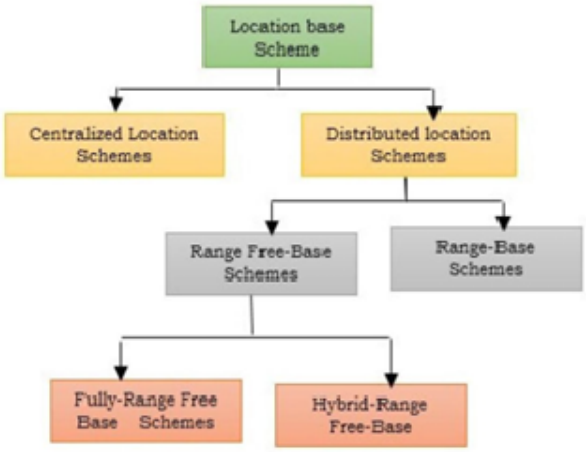


Fig. 3. Categorization of localisation schemes

is sent to all nodes in the network. The advantages of this approach removes all problem such as bandwidth, latency, energy-hole problem in every sensor node. The disadvantage of this scheme cause of available of global information which may be the case of security lost issues, the accuracy is less with respect to other algorithms.

2) Distributed localization technique: The decentralized schemes are divided into two types, free-range-based schemes and the range base schemes. Range free-based approach is based on using free-range measurement techniques for location estimation. Free range-based approach is divided into two sub types: one is fully free base technique and other one is hybrid free range technique. These classifications for unknown node consist on methods which are used for location estimation.

Range-Based Techniques: The range based methodologies are based on the physical characteristics of EM waves. Typically the following characteristics are used during range based position estimation, they are: Time Of Arrival (ToA), Angle of Arrival (AoA), Time Difference Of Arrival (TDoA), Received Signal Strength Indication (RSSI), Binary Sensors. Each of these have its own merit and difficulty during implementation.

ToA: Calculates the distances between nodes based on the time it took for an EM wave to travel between the two.

TDoA: It can be done in two ways. The first is to broadcast a signal from an unknown node to two anchor nodes and calculate the ToA in each. The second method is to implement a hardware solution which pulses two waves through the air with different propogation speeds.

AoA: Uses the angle by which the EM wave propogates over sensor nodes to calculate the distances between sensor nodes.

RSSI: This technique is based on the fact that signal strength of an EM wave decreases as it radiates outward in space. They are relatively cheap solution to implement positional estimation in a two dimensional space.

Free Range Base Localisation Techniques: The range free techniques are typically simple to implement. This is because RF techniques only rely on network connectivity and network topology information instead of physical measurements. Since no physical measurements are required, these methods can be easily used on constrained nodes. Some range free techniques include DV-Hop, MDS-MAP. In range free schemes, there is not special type of hardware used for distance estimation. Therefore, it has less costly and simple calculation for distance which is interesting area for researcher. In range-free technique also do not need to calculate the information of distance or information of the angle between unknown nodes by the consuming of energy for data collection, connectivity information and information calculation of the specific region. There are two sub categories of range-free schemes: one is fully-range free and second is hybrid-range-free techniques.

Fully-Range free localization schemes: In this type of schemes uses only one method based on connectivity and/or information of topology.

- Anchor Based Approaches
- Anchor Free Approaches

Anchor Based Approaches: The commonly fully-free range schemes depend on anchor sensor node which includes centroid, DV-Hop, APIT.

Anchor Free Approaches: The centralized free range algorithm is called MDS-MAP. Which is consist on MDS, it calculate the location of the unknown nodes which rely on the core information of closest sensor network. There is no anchor node in the network, this scheme is able to form relative maps that represent the current location information of sensor nodes.

Hybrid-Range free localization schemes: Hybrid-range-free scheme is combination of different methods based on topology and/or connectivity of location estimation of sensor nodes. These schemes are dividing into two categories, one is anchor base and other is anchor free approach.

IX. Security in SDN

Software-defined networking is a networking approach that aims to simplify network management and configuration. This happens by decoupling the control plane which encompasses all the network intelligence from the data plane which is the packet forwarding engine of the network. This essentially creates a global view of the network, allowing dynamic control and network programmability.Centralizing network intelligence results in access and control of the entire network through access to the controllers host server while the software used to control the network can be susceptible to bugs and other vulnerabilities.

The next threat in the data plane is attacking vulnerabilities in switches. Due to the nature of SDN, packet forwarding is already one of the problems presented , however by assuming control of a single switch, these packets can be dropped or

used to slow down the network, clone or divert traffic or inject forged requests and forged traffic to overload the network

X. Security Threats

Several security threats to SDNs have been identified .These security threats were then reviewed and further categorized by the functional layers of the SDN planes ,the application plane which houses SDN applications such as network management, the controller plane which contains the control logic framework and the data plane which contains the forwarding elements. Within the application plane, authentication and authorization is identified as one of the important challenges. The controller plane houses the control logic of the network and is therefore highly targeted by malicious users .Denial of service (DoS) attacks and data theft could be carried out through attacks on control plane communications. After gaining access to the controller, an attacker may launch distributed DoS (DDoS) attacks using a large number of switches under the control of an attacker.

XI. Solution

FRESCO allows security applications to be implemented on any OpenFlow controller . This enables security application development and policy implementation on NOX controllers . NOX defines a network operating system for OpenFlow controllers that allows management and centralized control application development. On the application layer, authentication and authorization is the main concern therefore most of the solutions to application layer security include validation and verification models . PermOF , a permission control application, is proposed to ensure that applications have controlled access to network resources

XII. Need for Wireless Sensor Network for Fire Detection and Alarm System

Traditional fire alarm system is based on a wiring network which have drawbacks and limitations such as inflexibility of the FDAS layout plan during building construction, and difficulties in renovation where the removal and relocation of traditional FDAS requires additional amount of work. To address these problems, a fire detection and alarm system that is based on wireless sensor network was developed .The FDAS is mainly composed of a fire detection node, a fire alarm node, and a fire alarm control panel.The wireless communication of the nodes was achieved using XBee as the wireless transceiver. The significant characteristic of XBee is the mesh routing protocol which helps the system to be more robust and flexible

XIII. System Overview

The overall system of the fire detection and alarm, as comprises of a LabVIEW-operated fire alarm control panel managing 2 zones.
Each zone comprises of eight detection nodes that has six smoke detectors, one heat detector, and one manual call point, along with one sounder and strobe light as the alarm node. All nodes that are in the same network automatically creates a mesh topology with the other neighboring nodes.
Lastly, the FACP is a PC workstation with LabVIEW software and a database. Data processing and monitoring is accessible in the software LabVIEW uses graphical programming language to create VIs (virtual instruments).
The algorithm of the FACP starts with the initialization of the LabVIEW software and creates the network through the XBee module which is serially connected to the PC workstation. Next, the coordinator node listens to the string of messages received from the different detection nodes. If the received string is a good alarm, a pop-up dialog box appears in LabVIEW before sending a control signal to trigger the notification device. At the same time, LabVIEW displays and stores the node address, zone of the node, date, and time of the alarm in the database.
Wireless Fire Alarm Node: The wireless fire alarm node is responsible for notifying individuals when there is a fire outbreak. The alarm node likewise consists of an XBee module, a microcontroller, and a CE listed notification device.
Fire Alarm Control Panel: The fire alarm control panel acts as the PAN coordinator of the entire system. The FACP manages the entire network from the detector nodes to the alarm nodes. At the same time,it is responsible for processing the data and sending the control signal to trigger the notification devices. The FACP primarily consist of a computer with a LabVIEW software installed and a ZigBee device.
Graphical User Interface: There are 2 main windows in the LabVIEW software: the front panel and the block diagram. The front panel highlights the “Admin page” of the GUI. This page signifies the access of authorized users where a username and password is required. Likewise, the setting the COM Port to be used where the XBee module has been inserted in the USB port and also the Baud Rate, typically “9600” or should be the same as the one configured in the XBee module.
Response Time: To determine the efficiency of the system from the time the detection nodes detects a large amount of smoke concentration or a rise of temperature, to the notification of the authorized user in the base station, and triggering the alarm nodes, different response time were acquired of thenodes deployed. The data gathered and analyzed serves as a primary reference in accordance to the standard discussed in National Fire Alarm and Signaling Code.

XIV. Conclusion

This paper is an overview of wireless sensor networks. It covers the area of deployment in WSN as well as the security and threats faced by Wireless Sensor Network. Excessive energy consumption is one of the major problems faced by WSN and various measures have been proposed for solving the same, among which one noticeable algorithm is the Energy Efficient Unequal Clustering algorithm (EEUC). Another major area in WSN where it is possible to encounter problems is in the deployment phase, which can resolved by various localization methods using software defined networking. As it is a connected network SDWSN face a lot of security issues

and threats. This paper proposes the method of FRESCO to overcome these issues.

REFERENCES

[1] J. N. AL-KARAKI, A. E. KAMAL. Routing techniques in wireless sensor networks: a survey. IEEE Wireless Communications, 2004, 4(11): 1-7.

[2] M. Chaw M. Thein, T. Thein. An Energy Efficient Cluster-head Selection for Wireless Sensor Networks. IEEE International Conference on Intelligent Systems Modelling and Simulation. 2010: 287-291.

[3] C. Jiang, W. Shi, X. Tang, et al. Energy-Balanced Unequal Clustering Routing Protocol for Wireless Sensor Networks(in Chinese). Journal of Software. 2012, 23(5): 1222-1232.

[4] A Review of wireless sensor network localisation based on software defined networking Olaus P. Cloete, Adnan M. Abu-Mahfouz, Gerhard P. Hancke, Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, South Africa, Modelling and Digital Science, Council for Scientific and Industrial Research, Pretoria, South Africa.

[5] Wireless sensor networks free-range base localization schemes Haroon Khan (Department of Computer Science, Bacha Khan University, Charsadda Pakistan), Muhammad Numan Hayat (Department of Computer Science, Abdul Wali Khan University,Mardan Pakistan), Zia Ur Rehman (Department of Computer Science, Bacha Khan University, Charsadda Paistan).

[6] H. I. Kobo, A. M. Abu-Mahfouz, and G. P. Hancke, "A survey on software-defined wireless sensor networks: Challenges and design requirements," IEEE Access, vol. 5, pp. 1872–1899, Feb. 2017.

[7] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Security in software defined networks: A survey," IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2317– 2346, Aug. 2015.

[8] K. Chelli, "Security issues in wireless sensor networks: Attacks and countermeasures," in Proc. World Congr. Eng., vol. 1, 2015.

[9] H. Modares, R. Salleh, and A. Moravejosharieh, "Overview of security issues in wireless sensor networks," in 2011 Third Int. Conf. Comput. Intell. Model. Simul., 2011, pp. 308–311.

[10] H. I. Kobo, A. M. Abu-Mahfouz, and G. P. Hancke, "A survey on software-defined wireless sensor networks: Challenges and design requirements," IEEE Access, vol. 5, pp. 1872–1899, 2017.

[11] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey,"Computer Networks, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.

[12] H. I. Kobo, A. M. Abu-Mahfouz, and G. P. Hancke, "A survey on software-defined wireless sensor networks: Challenges and design requirements," IEEE Access, vol. 5, pp. 1872–1899, Feb. 2017.

[13] M. Ndiaye, G. P. Hancke, and A. M. Abu-Mahfouz, "Software defined networking for improved wireless sensor network management: A survey," Sensors, vol. 17, no. 5:1031, pp. 1–32, 2017.

[14] S. Shin, P. Porras, V. Yegneswaran, and M. Fong, "FRESCO: Modular composable security services for software-defined networks," ISOC Network and Distributed System Security Symposium, vol. 2, pp. 1–16,Feb. 2013.

[15] Y. Ali and D. A. Akl, "Localization in wsn," 2013

Cross Platform for Mobile Application Development

Blessy Joseph, Janish P S and M R Aswathy
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Siji K B
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: siji.k.b@vidyaacademy.ac.in)

***Abstract*—Mobile applications, these days, are becoming significant in its usage as people tend to depend on it to perform several tasks using devices such as smartphones, tablets, etc. This increasing demand in the mobile application development has made it necessary to divide the platform into fragments that can support each mobile device. These smart devices could have different operating systems, different programming languages or even different development tools and this leads to a problem when someone wants to release an application in as many platforms as possible. Rewriting the application for each of these platforms turn out to be a burden in terms of high budget and development time. Therefore, it is important to overcome these problems and develop a cross-platform mobile application that will decrease the time to develop the application and also increase the number of users. As the app market expands, new challenges arise for developers. Creating useful applications is important in achieving extensive and cohesive use.**

***Index Terms*—Cross-platform, Mobile application development, Cross-Platform Applications, Native Applications, Hybrid Applications.**

I. Introduction

One way to bring a mobile application to multiple platforms is to develop a native application for each platform using the respective native SDK, such as Xcode for iOS, Android Studio for Android, etc. Native applications have full and direct access to the native platform functionalities and are capable of delivering the best performance and user experience on each platform. However, this approach affords minimum code reuse, and is the most expensive option in cost and time. The cross-platform mobile application development approaches allow a single code base to target multiple mobile platforms. The most significant advantages of these approaches are the cost and time savings to target multiple platforms. Mobile companies such as Apple, Google, Microsoft, etc., developed their own platform with specific tools. Along with the increasing demand of mobile application development it is also important to take care that the fragments of the operating systems support each platform. Model Driven Engineering (MDE), one of the emerging areas of Software Engineering has been proven useful in solving fragmentation, specially through defining apps recurring to domain models and Domain Specific Languages (DSLs). Several factors were considered to provide a better distinction between each tool, such as technology approach, supported platforms, development environment, type of the resulting app and device API support.

Fig. 1. Mobile App Development Process

II. Cross Platforms and Languages for Mobile Apps

Some cross platforms and languages for mobile application development are

A. Rhodes

Rhodes is a framework that manages enterprise mobile applications and data developed by RhoMobile. It uses the Model View Controller (MVC) architecture that separates the business logic (controllers) and data definitions (models) written in Ruby, from the visual interface (views), written in HTML, CSS and JavaScript. It features an app generator that creates controllers and views based on the models defined by the user.

B. PhoneGap

PhoneGap is an open source framework developed by Nitobi Software that allows the creation of mobile apps using HTML5, CSS3 and JavaScript. During the build process the web code is packaged together with the PhoneGap library into a native application execution shell. It provides a service called PhoneGap Build that allows developers to compile their PhoneGap apps in the cloud. Initially, the developers upload their app written in HTML, CSS and JavaScript and get back app-store ready application binaries.

C. DragonRAD

DragonRAD is developed by Seregon and it focuses on mobile enterprise applications. It is comprised of three major components: Designer, Host and Client. DragonRAD applications are compiled into Lua bytecode, which is executed by the interpreter in the DragonRAD runtime environment. Then, the application can be published to DragonRAD Host. The DragonRAD Client contains the code to interpret and run the application created in the Designer. DragonRAD comes with a default client that is loaded in the emulator or the device while debugging. However, a custom client can be build using the ClientBuilder service in DragonRAD's website when the application is finished.

D. Appcelerator Titanium

Appcelerator Titanium is a platform for developing mobile applications using web technologies developed by Appcelerator. Titanium applications are written in HTML, JavaScript and CSS. Titanium links JavaScript to native libraries, compiles it to bytecode and then the platform SDK (Android or iOS) builds the package for the desired target platform in order to run it. The output application contains mostly native code and rendering is executed natively. In addition, the output application contains a JavaScript interpreter runtime and a Webkit rendering engine.

E. Mobl

mobl is a free and open source domain specific language that uses HTML5, JavaScript and CSS and offers a language to built native-feeling web apps that can be deployed on several platforms. Its compiler automatically generates the related static JavaScript and HTML files ready to be deployed to any web. mobl allows the definition of user interfaces, data model, application logic and web service interfaces in a declarative manner. It is supported by Eclipse IDE providing features like error detection, syntax highlighting reference resolving or code completion.

F. Canappi mdsl

mdsl is a domain specific language developed by Canappi focused on describing an application in a technology and architecture independent way. mdsl allows the definition of a mobile application in a declarative way. The main concepts of this language are: Entity - represents the data model of the application. View - represents the data model of the application. Layout - represents a simple definition of a group of controls, such as buttons or text boxes, that are connected to views. Action - represents a standard action, such as "navigate to" another view. Connection - represents HTTP based operations which may accept a series of parameters and return a JSON or XML based response.

III. Factors Considered to Analyze and Compare the Tools

The factors considered to analyze and compare the tools were

1) Technology Approach: In this four classes of approaches were identified
   - Runtime: Runtimes uses several methods of code execution, such as interpretation, translation or virtualization. The typical approach consists in translating the code to bytecode and then, at runtime, that bytecode is executed by a virtual machine supported by the mobile device. Rhodes and Titanium belong to this category.
   - Web-to-native wrapper: Web-to-native wrapper is the approach used by PhoneGap and consists in a solution that packages web code with libraries inside a native app shell. Those libraries have the task of linking the web code to the native one.
   - App Factory: App Factories consist in visual tools that allow users to develop their app without code. Usually, they are based in templates and provide drag and drop features to generate the code. This approach is the one used by DragonRAD.
   - Domain Specific Language (DSL): DSLs are programming languages or executable specifications focused on a particular problem domain and that use abstractions and domain concepts to describe the apps. mobl and mdsl fall into this class through the application of concepts of the mobile domain.
2) Platform Support: PhoneGap and Rhodes are the tools that support more platforms whereas DragonRAD, mobl and mdsl support only two platforms.
3) Development Environment: PhoneGap, use a specific IDE to develop the applications. DragonRAD provides
4) Type of the Resulting App: Types of the app produced can be native, web or hybrid. Rhodes, Titanium, DragonRAD and mdsl generate native apps. PhoneGap generates hybrid apps while mobl generates essentially web apps.

IV. Developing Useful Mobile Applications in Cross Media Platforms

This presents a mobile application called "The Time Machine", offering location dependent information about the environment in a town. The app is evaluated through expert evaluation, as well as user evaluation conducted as a field observation with 15 participating users, measuring the level of usability in the application. Results are used to discuss guidelines on how to develop useful mobile applications. Knowledge about user behaviour, a clear application purpose,

an understandable interface, and cross-platform functionality are all important aspects for developers to consider.

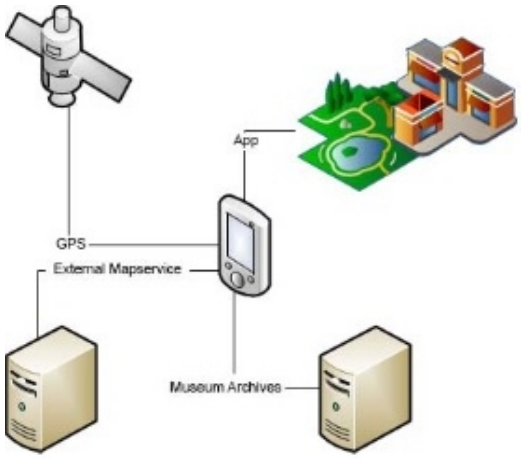


Fig. 2. A Protoype Application - The Time Machine

Our idea was to create a prototype application, "The Time Machine", offering citizens and visitors in a town location dependent information about e.g., houses, parks, streets from different ages, both past, current and future. The application communicates with both the town museum's digital archives and an external map service, as seen in Fig. 2. Application functionality includes geolocation-supported navigation; viewing information and photography's tied to the different locations; sharing information and photography's through social media; and a photo gallery.

Wireframes (see example in Fig. 3) were produced to display the user interface of the app and create navigations paths. The wireframes were used as a basis for an expert evaluation, where the lead interaction designer behind the official town application framework was interviewed regarding his view on our chosen application design. Expert evaluation results came to inform our final prototype design proposal as changes where made to both user interface and underlying ideas.

Fig. 3. A Wireframe

A. User testing

For our user evaluation, we chose a qualitative method. It is highly suitable when trying to gather in-depth data about the behaviour and opinions connected to the test situation. To examine the usefulness in an application, identification of errors in the use situation is often a good way to find weak spots in both concept and design. A heuristic study can help researchers to find interface problems and gather opinions from the user, hopefully resulting in better design. We therefore adopted a heuristic approach, as presented by Nielsen.

B. The usability aspect

Several ideas for improvement came up during both the user testing and the interview sessions. Many comments concerned The Time Machine interface. As for modalities, voices were raised for the inclusion of both augmented reality and voice-over commentary. Functionality was also addressed. Some wanted the possibility to "pause" the app when multi-tasking, but also to receive push notifications when walking past interesting sites. Addition of a compass for navigation was also desirable. Some users had comments about the visual content, e.g., photo meta data, output modalities (e.g., addition of video material), a time axis on the photo etc. The future photos where much appreciated and there were comments on extending this part even further in upcoming editions of The Time Machine. To be able to comment and interact with the persons responsible for city planning would encourage more people to use the app and trigger them to use it more than once. Some users also suggested the addition of guided history tours, or up to date information about events taking place here and now.

C. The information aspect

Many saw the application as a "seasonal app", best used during summer or holidays, when there is much spare time and the weather is fine. Constantly updating The Time Machine might be an important building brick in trying to achieve cohesive use.

D. The design and navigation aspects

Users appreciated the simplicity of the map and found it easy to navigate. There where some interface problems regarding markers that where not clickable and also some confusion about the function of the photo archive included. This was mainly due to a lack of content in the application prototype regarding this particular feature. A clean interface was found helpful during navigation. Easy access to the search function was also stressed as an important thing for increasing usefulness. The small display of the smartphone used in the test however proved to be challenging when important buttons were placed too close to each other. One test subject was in a wheelchair, and she preferred navigation through clicking rather than swiping.

E. Junk apps

During the interview sessions, test subjects had the opportunity to give their general opinions of what characterizes both good and bad apps. It is clear that users can be quite judging, and that far from all applications turn out to be useful. Some users had downloaded apps that automatically closed down

other apps, no longer needed. This pinpoints a problem that users have too many apps that become too cumbersome to control.

V. A Performance Evaluation of Cross-Platform Mobile Application Development Approaches

The different applications that can be developed using cross platforms are:

A. Packaged Web Apps/Hybrid Apps

A popular and economical approach for cross-platform mobile applications is to develop mobile applications using the standard web technologies (HTML,CSS, and JavaScript) which are supported by all mobile platforms. For this category, applications developed using Apache Cordova are included in our study.

B. Proxy Native Apps

Appcelerator Titanium is a JavaScript based framework and tool suite aims to remediate some of the shortcomings of the packaged web app approach by designing a unified and higher level, i.e., more abstract, JavaScript API for building native UI. Implementations of the unified JavaScript API are provided for several popular mobile platforms, including Android and iOS. Applications built using this approach can look and behave exactly as native apps.

C. Cross-Compilation with Native API Binding

Another approach is to define API bindings to allow programs in one language to consume the libraries of a different language, and to cross-compile application source code to binaries of different target platforms. Xamarin is an example of this approach, which allows mobile applications to be developed in C# and then compiled into binaries for several mobile platforms.

1) Design of the experiment: A set of mobile apps with identical features using different SDK's and tools were developed to measure their performance characteristics on the same devices under identical conditions, we have designed three benchmark mobile apps, each focusing on a different aspect of app performance. Each of the benchmark apps consists of two screens: the initial screen is a configuration screen for setting the parameters of the benchmark app. The second screen is the main screen which displays a screenful of contents of variable sizes. The size of the contents is controlled by one or more parameters of the benchmark app.

VI. The Impact Of Cross-Platform Development Approaches for Mobile Applications from The User's Perspective

It is imperative for developers to be aware of the impact of different cross-platform development approaches on the quality of a mobile application. The quality of an app, more accurately, the perceived quality of an app may be assessed from reviews posted by real users of the app-on-app marketplaces such as Google Play and Apple App Store. User reviews have been regarded as a reliable source in assessing the perceived quality of apps. Here, reviews are categorized in to four-performance, usability, security and reliability. First, we discussed about the methodology used in building the dataset of user reviews from both Google Play and Apple App Store for a set of 50 apps. The classified user reviews are analysed using a metric that quantifies the proportion of reviews in which users complain about a particular quality concern.

$$Complaintdensity(qc) = (ReviewswithqcComplaint/ TotalReviews)\%$$

where qc=quality concern (i.e., performance, usability, security, or reliability.

The next is App selection. We need to identify various tools that support the most important cross-platform development approaches in practice today.

We created initial models by skimming through some of the reviews to identify a preliminary set of keywords pertaining to complaints in each of the four quality concerns. The initial models were then used to classify the entire data set of reviews. We then created a training set by random selecting 50 reviews from each of the four groups (i.e., performance, usability, security, and reliability complaints) of reviews as classified by the initial models. The training set, therefore, contained 200 reviews in total. The purpose of the training set was to assist in the iterative validation and refinement of the model. We analysed the data set of user reviews in two different ways. In RQ1, we compared the different development approaches and the tools that support them. In RQ2, we analysed a single app that underwent a change in its development approach from cross-platform hybrid to native. In the subsections that follow, we present the results from our analyses.

We created a model to classify users reviews in performance, usability, security, and reliability complaints. We evaluated these models using Cohen's k, precision, and recall. The usability model was the worst with k = 49.26%, however, performance, security, and reliability models had k values of 65.78%, 72.75%, and 92.80%, respectively. Perhaps, more complex NLP techniques could improve the usability model. The key results from our work are as follows:

- Android users had a more negative perception about the performance, reliability, and usability apps developed with hybrid approaches than with the traditional native approach.
- Security was the only quality concern in which the hybrid approach outperformed interpreted/generated in Android apps.
- The results from iOS apps were less consistent. Native apps had a higher complaint density than hybrid apps across all quality concerns except performance. Interpreted/generated apps were better than native only in terms of usability and performance.
- When Facebook changed from hybrid approach to native development, the density of user complaints along performance and reliability reduced substantially (reduction of 2.72 pp and 13.45 pp, respectively).

- PhoneGap and Unity3d had the highest performance complaint density in iOS apps.
- PhoneGap had the lowest performance complaint density in Android apps and the lowest reliability complain density in iOS apps.
- PhoneGap Android apps, native iOS apps, Kony, Alpha Anywhere, and Appcelerator had the highest reliability complaint density (greater than 15%).
- Coco2d iOS apps had the highest usability complaint density (2.28%). Apps developed with Corona SDK, Alpha Anywhere, and Kony had the highest security complaint density(0.73%, 0.71% and 0.66%, respectively).

VII. Challenges for Mobile Application Development

The technical analysis of major operating systems and development platforms to identify the strengths and weaknesses of these platforms.

1) Apple/iPhone: With its large screen, good hardware/software capabilities like 2G/3G and Wi-Fi network connectivity, camera, motion sensors, touch display, user friendly interface to the World Wide Web etc., has been a huge success and a trend leader in the smartphone market. When iPhone was launched it was different from the way smartphones were designed. Available smartphones at that time where extension of PC and targeted towards business people. It uses iTunes Store, App Store and MobileMe. It comes with a YouTube client preinstalled and the newest version support HD video playback upto 720p.We have major difference between iPhone and other platforms like Vodafone live and iMode. The primary revenue source for Vodafone and NTT DOCOMO are the generated traffic as none of them are handset companies. In the beginning Apple also got shares of the generated traffic, this was due to Apples marketing strategy. Another specificity of iPhone is that is merely a platform rather than a sophisticated handset, is the launch for 'App Store' in 2008.Here Apple has created a market place where the application providers can meet the consumers and sell their application directly to the consumers. The model is then that Apple gets a share of revenue, with a 70/30 split, with 70% of the revenue to the developer and 30% to Apple. In App Store, it is possible to distribute free Applications but developers must pay 99U$ per year.

2) Google/Android: Google acquired the small mobile software mobile software developer company, Android, Inc.in July 2005.It also include the mobile market. The aim has been to develop a common open-source operating system that any mobile device can run on and to make developing applications for mobile phones more general, and hardware independent. The development of Android is taking place in the framework of 'The Open Handset Alliance'(OHA) with a number of major companies from the mobile industry. Currently, the Android market charges 25 USD to be a registered developer. As well, a developer can purchase an Android Developer's Phone for almost 400 USD. An important difference between Android and iPhone platforms is that Android is not tied to specific hardware and offers drivers for many of the common hardware components on the mobile phone.

3) Other platforms:

- Nokia/OVI store
  Nokia has also created its own App Store called OVI store. It is new which means it contains limited amount of applications. The applications are not as creative and advanced in functionality and graphics. This is important to follow in future.
- Blackberry
  Its App Store is 'App World'. The applications in this store are more traditional like getting world weather, access to existing web services with less focus on the user experience.
- Windows mobile
  Microsoft launched access to a new application store called Windows marketplace. The amount of application is limited and the functionality is not creative.
- Palm
  Palm's application store is called 'Palm App Catalog' and currently there are only a few applications available.

VIII. A Comparative Study Of Cross-Platform Mobile Application Development

For comparison, a sample practical project, a coin wallet application based on block-chain was developed. The sample application was implemented in Android native and three different cross-platform frameworks respectively.

The Coin wallet is a financial application running on the phone that allows users to transfer their money to others. The project contains a server-side and four client-sides. The server-side is a daemon service running on a workstation to transmit the blocks of data to clients. The client-sides include four applications that were implemented by Android native, React Native, Flutter, and Xamarin respectively. Each client application is composed of the UI, wallet, encryption algorithm, and network implementation. The UI module is in charge of screen rendering related functions by popping up a notification, accepting input text, and presenting rich graphic elements on the phone. The wallet is responsible for transferring or collecting money from others. The encryption module needs to generate, store, and compute the key on the phone. The network module encapsulates RESTful API to make the communication between clients be easy.

Consider an example of workflow for transferring money of Rs. 30 from Client2 to Client1 using the sample project. First of all, both clients need to input the server's IP address in order to ensure all applications communicate through the server properly. Secondly, the receiver (Client1) needs to be ready by clicking the "Collect" button to show its account (QR code). Thirdly, the sender (Client2) is ready of transaction by clicking the "Transfer" button; entering the amount (e.g. Rs. 30); and scanning the QR code on receiver (Client1) to transfer the money. Fourthly, the receiver (Client1) receives the money

and the balance is updated on both clients. Finally, both clients click "Transaction" button to check their transaction records.

A. Research Findings

1) Workload: Android Native takes the highest workload in code lines than the other three cross-platform frameworks. Flutter, using Dart, requires the lowest workload that is only 47% of the total workload from Android Native. Xamarin produces a higher workload than React Native.

2) Performance: The Android Native development takes the lowest CPU usage, the smallest memory usage, and the most compact binary package. Also, the study found that Xamarin runs more efficient not only in CPU usage, but also in memory usage with the smallest binary, while React Native has the lowest performance with the biggest binary package.

3) UI development: Android Studio is an integrated development environment that covers all the life cycle of application development including UI designing, debugging, and deployment. React Native divides the processes into several parts that are developed by different single tools. It needs to employ a terminal or command line interface in order to install the environment and also requires to apply different editors in order to edit the code. Moreover, both its debugging and publishing application have to go through the command line. In a word, the overall use of React Native is not as user friendly as Android Studio. Flutter is more user-friendly than React Native in the sample project development experience. It utilises Android Studio for code editing and debugging, and for project deployment and also allows developers to preview the change when debugging.

Xamarin is supported by Visual Studio that provides many features to help developers in their code editing and debugging, and products delivering. Visual Studio also has features that allow developers to distribute their applications directly to Google Play and Apple Store. In a word, among the above three cross-platforms, Xamarin offers the best user's experience in UI development for the sample project.

IX. Conclusion

In this paper, we presented a study on the different cross-platform tools for mobile application development. Different cross-platform technologies are appropriate to different development. Firstly, the Android Native framework is preferred if the project is sensitive to running speed and memory usage. Secondly, React Native is rather suitable for web developers because JavaScript is widely used in web programming. Thirdly, Flutter is recommended for those who are familiar with C++ or Java or who need to port current projects from a native framework to a cross-platform framework because it employs a C-like programming language, Dart. And finally, Xamarin is more suitable for .NET developers or those who are concerned about performance because it is easy to cross-platform development and to implement the most efficient application among cross-platform frameworks using C# in Visual Studio.

References

[1] André Ribeiro, "Survey on Cross-Platforms and Languages for Mobile Apps" 2012 Eighth International Conference on the Quality of Information and Communications Technology.

[2] Dongliang You and Minjie Hu, "A Comparative Study of Cross-platform Mobile Application Development", 12th Annual Conference of Computing and Information Technology Research and Education New Zealand (CITRENZ 2021).

[3] Allen Hammershoj, "challenges for mobile application development", Lautrupvang 15, 2750 Ballerup, Copenhagen, Denmark.

[4] A.Charland and B.Leroux, "Mobile Application Development. Web vs Native", Commun.ACM,54(5):49-53,May 2011

[5] Xiaoping Jia, Aline Ebone and Yongshan Tan, "A Performance Evaluation of Cross-Platform Mobile Application Development Approaches", 2018 ACM/IEEE 5th International Conference on Mobile Software Engineering and Systems.

[6] Cajsa-Tora Hermansson, Mikael S¨oderstr¨om and Dan Johansson, "Developing useful mobile applications in cross media platforms", 2014 Eighth International Conference on Innovtive Mobile and Internet Services in Ubiquitous Computing.

[7] A. Charland and B. LeRoux, "Mobile Application Development: Web vs. Native", Commun. ACM, vol. 54, no. 5, pp. 49-53, New York, May 2011.

[8] Vision Mobile, "Cross-Platform Developer Tools 2012: Bridging the worlds of mobile apps and the web ", February 2012.

[9] M. Book, S. Beydeda and V. Gruhn, "Model-driven Software Development", Springer, 2005.

Metaverse: The Future of Internet

Anjitha N A, Athul P S and Catherine Joshi
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Salkala K S
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: salkala@vidyaacademy.ac.in)

***Abstract*—The Metaverse is the post-reality universe, a perpetual and persistent multiuser environment merging physical reality with digital virtuality. It is based on the convergence of technologies that enable multisensory interactions with virtual environments, digital objects and people such as virtual reality (VR) and augmented reality (AR). Hence, the Metaverse is an interconnected web of social, networked immersive environments in persistent multiuser platforms. It enables seamless embodied user communication in real-time and dynamic interactions with digital artifacts. This paper summarizes the development status of the Metaverse, starting from the definition of the Metaverse, how metaverse helps in improving the education and gaming and about the challenges of applying metaverse.**

***Index Terms*—Metaverse, Avatar, Ubiquity, Meta-societies, Digital Currency, Economy System, Temporal demand**

I. Introduction

The word Metaverse is a portmanteau of the prefix "meta" (meaning "beyond") and the suffix "verse" (shorthand for "universe"). Thus, it literally means a universe beyond the physical world. Metaverse is a new type of Internet application and social form that integrates a variety of new technologies. It provides an immersive experience based on augmented reality technology, creates a mirror image of the real world based on digital twin technology, builds an economic system based on blockchain technology, and tightly integrates the virtual world and the real world into the economic system, the social system, and the identity system, allowing each user to produce content and edit the world. Metaverse is still a concept that is constantly evolving, and different participants are enriching its meaning in their own ways. From the user's point of view, the demands on the freedom of the virtual world, the content and interaction methods of the Internet, etc. are constantly increasing.

II. Current Status and Future Possibilities of Metaverse

The Metaverse refers to a fully immersive three-dimensional digital environment in contrast to the more inclusive concept of cyberspace that reflects the totality of shared online space across all dimensions of representation. Virtual world development now faces a major new challenge: how to move from a set of sophisticated but completely independent immersive environments to a massive integrated network of 3D virtual worlds or Metaverse, thus establishing a parallel context for human interaction and culture. The current work has defined

Fig. 1. Metaverse

success in this effort as achieving significant progress with regard to four features that are considered central elements of a fully-realized Metaverse: realism, ubiquity, interoperability and scalability.

A. Realism

Technologies and techniques for improving immersive realism, especially for the senses of sight, hearing, and touch, continue to evolve independently of virtual worlds due to their applications in other domains, such as games, entertainment, and social media. Bringing these developments into the Metaverse is typically a phase behind innovations in these other contexts due to the additional requirement for bidirectional information flow and real-time performance. Moving forward, realism-related work must achieve either concurrent input/output capability or real-time performance, or both, in order to achieve a useful role in virtual world applications.

B. Ubiquity

Progress in ubiquitous availability and access to virtual worlds may very well take care of itself, as the increasing power, diversity, and availability of general computing devices as well as the capabilities of browser-based implementations continue to advance. As long as virtual world developments move alongside general ubicomp developments, moving in and

out of the Metaverse may become as convenient and fluid as browsing the Worldwide Web is today. In contrast, achieving ubiquitous identity will require more concerted and coordinated effort, both technologically and socially/politically, to allow one's information to flow more freely among various milieus. Social networks are beginning to move in this direction though have not fully reached it yet, and virtual worlds must follow suit in this area. It is also possible that virtual worlds may take a leadership position in this regard because virtual world artifacts may be more closely linked to one's digital persona due to the immersive environment.

C. Interoperability

Interoperability in virtual worlds currently exists as a loosely connected collection of information, format, and data standards, most of which focus on the transfer of 3D models/objects across virtual world environments. Although many standards have been proposed, no single one has gained long-term traction. However, virtual world interoperability is not solely limited to 3D object transfer: true interoperability also involves communication protocol, locator, identity, and currency standards. The wide-ranging requirements and scope of digital assets involved in virtual worlds have the potential of making the Metaverse the "killer app" that finally leads the charge toward seamless interoperability.

D. Scalability

Virtual world technologies are currently in an initial stage of departure from highly centralized system architectures, with the distribution of load being pre-dominantly focused on contiguous regions or areas. An overall architectural strategy based on DSG, peer-to-peer openness as seen in Open Cobalt and Solipsis for broad distribution of load, and reduction/optimization techniques like action-based protocols appears to hold the most promise for massive scaling of an integrated virtual world system. Progress in virtual world scalability implies progress in the scalability of many other types of multi user multitiered systems.

III. Enhanced Playability for Mobile Gamers

Virtual environments are an implementation of the metaverse concept, a simulated world populated by a massive number of synthetic avatars that are controlled via Internet. The avatar is a tool allowing humans to interact with a metaverse. Over the last years, the interest of the gaming industry has led to the implementation of many metaverses called virtual worlds (VWs). Many of them are so realistic that they have an economy, government, and currency of their own. Due to the unreliable nature of the wireless medium and to the mobility, this kind of gamers would require special mechanisms to maintain a good level of playability, at least from a technological point of view. The mobile nature of wireless gamers can be often the cause of interruptions and lags in the communications between devices and gaming infrastructure. We propose the playing session (PS) mechanism to enhance the gaming playability while maintaining good levels of equity and fairness between all users.

The combined effect of a widespread availability of Internet and the development of a new generation of wireless devices will lead to a massive amount of gamers interested in VW. when a mobile device connects to an AP that belongs to a different internet service provider (ISP), the mobile device should be able to obtain a new valid IP address. This could lead to an incorrect management of the network communications, and even, in a positive case, to the reconfiguration, while resume mechanisms might require many seconds to bring the system back to a working state. In this scenario, users playing from a mobile device and therefore using unreliable networks could be severely disadvantaged with respect to "wired gamers." They might lose some match turns and be also disconnected from the whole system. In all of these cases, the overall playability would be dramatically decreased.
We define playability as the user's satisfaction while playing. In this scope, we designed and implemented a new mechanism called playing session (PS) that is in charge of monitoring the client-server communication channel. The mechanism will be triggered by network failures, mainly due to the gamers' movement. The PS is composed by two parts: the participatory framework (PF), in charge of detecting and reacting to the network failures and the mimicking mechanism (MM) that tries to reproduce, with an adequate level of mimesis, the user behaviour. The PF is present on both sides (client and server). On the mobile device, the PF is able to detect if the currently used game server is unreachable or reachable with a certain delay. On the game server, the PF is able to detect whether the application on the mobile device is experiencing problems. In this case, the PF gives the control of the player's avatar to the mimicking mechanism (MM) that will be in charge of playing, until the communication channel is restored. In this way, the game can continue the progress avoiding that the whole system is affected by the fault of a single player. The reason for implementing metaverse in gaming are:

A. Interactivity:

The gamers expect to play fluently, without taking care of problems deriving from devices' limitations or due to the unreliability of communications. The idea is to relax the temporal constraints of the game progress up to the sensorial perceptivity threshold. This means that the length of a single turn is upper bounded by this threshold. In this way, it is possible to use all the available time to wait for "delayed moves".

B. Coherence:

It is really unpleasant for gamers to find frozen avatars in the middle of the scene. A common and really unsatisfactory solution to this problem is to set some prefixed actions to support an "orphan avatar". In order to avoid this unpleasant experience, it is important to maintain a good level of coherence inside the game. In this case, we define coherence as uniformity among players: each player should be able

to participate in the game in the most correct way, without limitations severely degrading his experience. The real world is very different: network faults are frequent and the usability of the input devices is often unsatisfactory. In this scenario, an MM would be able to increase the level of the game coherence.

C. Equity/fairness:

Interactivity and coherence are the bases for achieving a good level of equity/fairness in the game. In gaming, the aim is to guarantee substantial equity among all players. The problems due to network communications can have a high impact on the game equity. From our point of view, avatars and gamers should be decoupled, we consider avatars as processes that, in case of network failures, are separate from gamers. The chance for each gamer to play the same number of actions with the same frequency in the match is a key point to evaluate the game equity, and therefore both aspects should be carefully measured. The PS mechanism gives all gamers the same chance to win the game.

PF is composed by a couple of modules, the user participatory framework (UPF) and the avatar participatory framework (APF). The UPF, which is accommodated in the gamer device, checks the state of communications and verifies if the gaming architecture is reachable or not. On the server side, APF monitors the communications with gaming devices. From a technical viewpoint, the current implementation of the PS mechanism works in the client-to-server side of the gaming architecture. As a future evolution, the mechanism could be also extended to the server-to-server side: in this case, aiming at reducing the impact of network failures in the communications among servers.

IV. Fusing Blockchain and AI With Metaverse

Metaverse as the latest buzzword has attracted great attention from both industry and academia. Metaverse seamlessly integrates the real world with the virtual world and allows avatars to carry out rich activities including creation, display, entertainment, social networking, and trading. Thus, it is promising to build an exciting digital world and to transform a better physical world through exploration of the metaverse. Artificial intelligence and blockchain technologies are expected to play essential roles in the ever-expanding metaverse. For example, metaverse uses Artificial intelligence and blockchain to create a digital virtual world where anyone can safely and freely engage in social and economic activities that transcend the limits of the real world. Exploiting metaverse, the application of these latest AI and blockchain technologies will be accelerated as well.

A. Artificial Intelligence in Metaverse

Artificial Intelligence is a research discipline proceeded on the basis of the hypothesis that every aspect of learning can in principle be so precisely described. The state-of-the-art AI studies focus on machine learning, deep learning, and reinforcement learning in the fields including computer vision, decision-making, natural language process (NLP). Intuitively, breakthroughs of artificial intelligence in the real-world people to use it to realize the metaverse. AI-driven non player characters (NPCs) are computer-operated characters who act as enemies, partners, and support characters to provide challenges, offer assistance, and support the storyline. While from the game perspective, most of the human-looking NPCs are not intelligent enough to interact with players in certain game genres which requires strong tactical decision-making abilities.

B. Blockchain in Metaverse

Blockchain is widely believed as one of the fundamental infrastructures of metaverse. Blockchain is expected to bridge isolated small sectors together and provides a stable economic system, which helps offer transparent, open, efficient, and reliable rules for metaverse. For example, Hash algorithms and timestamp technologies provide metaverse users the traceability and confidentiality of underlying data. The conventional Blockchain architecture includes data layer, network layer, consensus layer, incentive layer, contract layer, and application layer. Cryptocurrency is one of the main applications under spotlight empowered by blockchain. It also makes blockchain more popular. The trust of a wide range of users supports the value system of cryptocurrencies, and drives both the circulation and trading of cryptocurrencies. The price of mainstream cryptocurrencies keeps hitting their record highs.

Currently, the economic activities in metaverse mainly include the auction of virtual assets, including land, scarce items, precious real estate, the development and leasing of land, the rewards of game tasks, and the profits from investing cryptocurrency. Thus, metaverse invokes a new form of funding that draws inspiration from both the real and virtual world. Different from the physical world, the digital creation in virtual world might be unlimited. The identity of digital objects determines value instead of the undifferentiated labour in conventional economy. In the field of digital creation, it is necessary to develop authoring tools to enable the users to produce original content easily and gain rewards efficiently with low cost. Those tools could improve the enthusiasm of content producers of metaverse.

In future metaverse, people prefer to turn to their virtual cabinet to select digital outfit, while companies begin to hype the virtual skins, virtual clothing, even virtual estates with high price which will block a large portion of players to join in metaverse. Hence, it is necessary to propose particular governance mechanisms under the cooperation of worldwide companies. Furthermore, how to establish a digital currency system that enables the currency exchange between metaverse and the physical world still remains an open issue. Currently, the concept of metaverse is mainly used and propagated by companies such as Roblox and Meta (Facebook previously). It is predictable that the most popular ecosystems in the near future are built and operated by these giant corporations. Tiny metaverse has only a few application scenarios. In contrast, the macro metaverse would include all scenarios required by users.

To realize such a vision, large companies need to cooperate to create a huge unified metaverse.

V. Higher Education in The Metaverse

Metaverse has a significant role to play in higher education. Business and industry will be looking for an educated workforce ready to meet the challenges these new environments present, from new modes of marketing, design, and manufacturing to new kinds of management and organizational leadership models that leverage virtual environments as one option in a suite of online tools serving the needs of companies and their customers. As virtual economic activity and the delivery of virtual goods and services increase, so too will the need for new areas of research and scholarship. The social sciences can help us better understand how human behaviour in these environments is similar to or different from behaviour in the real world, while computer scientists and engineers will be needed to increase capacity, usability, safety, and crossover to real-world facilities and objects. Businesses are already seeking interns with experience in virtual worlds and are partnering with researchers to develop innovative uses for existing platforms. This trend is likely to increase as virtual worlds gain broader adoption and more widespread use in the commercial sector.

Higher education institutions can benefit from many of the same cost-saving techniques as businesses by leveraging the collaborative capabilities of virtual worlds for conducting the "business" of education. From enabling internal collaboration and planning meetings to providing student services on a virtual campus, virtual worlds hold enormous promise for providing a platform for faculty, staff, and students to interact in an environment that can be entirely flexible to accommodate different needs. The course/learning management systems like Blackboard would be quick to explore this market, but we've seen few examples of this to date. Until we see greater adoption by higher education institutions and more demand from faculty and students, demand that will trickle up to those making the purchasing decisions, service providers are unlikely to begin to interface with virtual worlds platforms. Because most virtual worlds platforms are proprietary and the underlying technical standards that help different systems talk to each other have not yet been developed, existing learning management systems would have to start from scratch in a new market.

Finally, higher education institutions facing a plethora of choices offered by virtual worlds will need to carefully consider a number of strategic and legal issues. In recent years, the trend toward "software as a service" has been growing: instead of purchasing a software product that is installed and hosted on the institution's own servers and managed by internal employees, the institution pays an outside company to host the software application and provide service to customers via the Internet. Virtual worlds platforms largely seem to be following this model, which places higher education institutions in the position of relying on the service provider for everything from data security, uptime, and reliability to governance of the legal relationships between the institution and its faculty, staff, and students under the provider's terms of service.

VI. Metasocities in Metaverse

The topic Meta-Societies in Metaverse, which runs parallel with real societies, and expands the living and working space for humans through the virtual real interactions and closed loop feedback with real societies, Meta-Societies can provide description, prediction and prescription for the real societies.

A. Meta-Enterprises and Meta-Cities

The artificial society, computing experiment, parallel execution theory and cyber physical social system technology provides a theoretical and technical foundation for the development of Meta-Enterprises and Meta-Cities. These are considered as mapping of real enterprises and cities in virtual cyberspace. The virtual enterprises and cities are running parallel to real enterprises and cities. Virtual elements in Meta Enterprises and Meta Cities can be used to analysis the decision making with computational experiments. Meta Enterprises and Meta Cities can improve the decision and operation efficiency of real enterprises and cities. Meta Enterprises and Meta Cities are consistent with real enterprises and cities in terms of scale, behaviour, and system characteristics, and can provide an effective means to realize the "multiple world" view. Before any major decisions or operations in real enterprises and cities, numerous computational experiments should be first conducted in Meta Enterprises and Meta Cities, to analyse, evaluate, and optimize the decision-making processes as well as the master skills and resources required for a successful real execution in the shortest time, with the least energy and cost. This will greatly improve the efficiency and effectiveness of the decisions of real enterprises and cities.

B. Meta Economics for Meta Enterprises and Meta Cities

The Meta Economic system can be regarded as the virtual economic system corresponding to the real economic system. With the computational economic experiment approach, small economic data in the real economic system can be converted into large virtual economic data in the Meta Economic system, and then artificial intelligence algorithms such as machine learning, deep learning, and reinforcement learning can be used to transform virtual economy big data into deep intelligence. In the Meta Economic system of Meta Enterprises and Meta Cities, there are many software-defined entities and virtual human resources such as knowledge robots or advanced digital assistants, to describe, predict, and prescribe the operation of the real economic system. The Meta Economic system can provide a detailed description of the knowledge needed in the real economic system, as well as economic theories, procedures, processes, and related operations. As such, it can be used by managers of real enterprises and cities for the purpose of learning and training.

C. Meta Management for Meta Enterprises and Meta Cities

Meta Management can realize the management of real enterprises and cities through the management of Meta Enterprises and Meta Cities. In Meta Enterprises and Meta Cities, the organizational structures are organized with the form of blockchain-based distributed autonomous organization (DAO), which has the characteristics of distributed and decentralized, autonomous and automated, and organized and ordered. Therefore, the Meta Management for Meta Enterprises and Meta Cities is mainly the management of DAOs corresponding to Meta Enterprises and Meta Cities. According to the management decisions of real enterprises and cities, the corresponding Meta Management rules of Meta Enterprises and Meta Cities are constructed. The Meta Management rules are automatically implemented through smart contracts in Meta Enterprises and Meta Cities, and according to the operation results, the Meta Management rules are continuously adjusted, so as to achieve the expected Meta Management effect. Finally, the adjusted Meta Management rules are used to optimize the real management decisions.

D. Meta Societies in Metaverse: Toward parallel Societies

The Meta Societies running in the virtual cyberspace can effectively break the space, time, and economic restrictions of humans, and create a new way of life, work, and communication for humans in real societies. As the mapping of real societies in virtual space, Meta Societies can verify and evaluate the decisions in real societies using computational experimental approaches in Meta Societies, and prescribe various decision-making in real societies through virtual-real interaction, closed-loop feedback, and parallel implementation of the two societies. In the future parallel societies, the Meta Societies and the real societies will run in parallel. The changes of the real societies will affect the Meta Societies, and the changes of the Meta Societies will also bring the changes to the real societies.

VII. The Challenges of Entering The Metaverse

Information technologies exist to enable us to either do things we have not done before or do familiar things more efficiently. Metaverse enables novel forms of engrossing telepresence, but it also may make tasks more effortless. Such technologies increasingly facilitate our work, education, healthcare, consumption and entertainment; however, at the same time, metaverse bring a host of challenges. In the current study we conducted a 2 (AR: with vs. without) × 2 (VR: with vs. without) between-subject experiment where participants faced a shopping-related task to examine a proposed series of hypotheses. The NASA Task Load Index (NASA-TLX) was used to measure subjective workload when using an XR-mediated information system including six sub-dimensions of frustration, performance, effort, physical, mental, and temporal demand. The findings indicate that AR was significantly associated with overall workload, especially mental demand and effort, while VR had no significant effect on any workload subdimensions. There was a significant interaction effect between AR and VR on physical demand, effort, and overall workload. The results imply that the resources and cost of operating XR-mediated realities are different and higher than physical reality. In the field of information systems, individual workload or the lack thereof is often operationalized simply as the general ease of use of an information system in a given task according to the technology acceptance model (TAM). The current study assumes that both AR and VR technology will lead to a high overall workload compared with nontechnology mediated reality, but will influence different subdimensions of workload. In addition to mental demand, prior studies postulate that using XR technologies can result in higher degrees of physical exertion. Temporal demand is related to perceived time pressure, which is the perception that there is an inadequate amount of time available to carry out the given tasks. Based on the assumption that XR can generate both higher mental and physical demands, and considering the notion that perceived effort is caused by the level of work an individual has to expend in order to reach a certain level of performance.

With the development of information technologies, human computer interfaces and computing power, our daily lives are increasingly lead in technology-mediated environments. The present study investigated how XR (AR and VR) influence the six dimensions of workload and overall workload based on a 2 by 2 between-subject experiment in the retail context. The results from granular analysis indicate that AR was significantly associated with overall workload, especially mental demand and effort, while VR had no significant effect on any of the dimensions of workload. Additionally, the results of the interaction effects show that a combination of AR and VR compared to a single technology would not increase the difficultly of completing tasks, and may even decrease the difficulty. By filling the current gap in research on XR of information systems, human computer interaction and management science, the current study made a considerable research contribution as well as practical guidance for XR designers, developers and practitioners.

The analysis of workload in the present study deepens the understanding of what factors affect usability and efficiency, and what factors are important to users in extended realities, contributing to the corpus of information system and technology-oriented information science. There is a lack of comprehensive understanding and accurate measurement relating to system efficiency and usage barriers in the extant XR-related literature. The study findings have valuable implications, mainly for XR designers and programmers, and also for practitioners from different areas, especially those in the retail sector. Due to a lack of theoretical guidance and granular analysis of perceived workload, many program developers and designers face the challenge of finding good solutions of how to improve the perceived usability, ease of use, usefulness and efficiency extended towards XR technologies and systems. The shopping environment was carefully considered and designed for investigating the matter of perceived workload related to XR, and selected as the research context in the current study. Shopping tasks usually require a lot of physical and cognitive

resources, and have been used in many XR-related studies. A second limitation of the research design in our study is the lack of multisensory experience in virtual reality, which is associated with user experience.

VIII. Conclusion

The Metaverse is not a new concept. However, in the context of MR, it can bridge the connectivity of social media with the unique affordances of VR and AR immersive technologies. If the interplay among them is unleashed creatively, it promises to transform many industry sectors, among them distance online education. New models of Meta-education, Metaverse-powered online distance education, can emerge to allow rich, hybrid formal and informal learning experiences in online 3D virtual campuses. Technology giants such as Apple and Google have ambitious plans for materialising the metaverse. With the engagement of emerging technologies and the progressive development and refinement of the ecosystem, our virtual worlds will look radically different in the upcoming years. Now, our digitised future is going to be more interactive, more alive, more embodied and more multimedia, due to the existence of powerful computing devices and intelligent wearables. However, there exist still many challenges to be overcome before the metaverse become integrated into the physical world and out everyday life. We call for a holistic approach to build the metaverse, as we consider the metaverse will occur as another enormous entity in parallel to our physical reality.

References

[1] Nick Lambert, *Beyond nfts: A possible future for digital art*, ITNOW, 63(3):8–10, 2021

[2] KR Malik and M Farhan,*Merging virtual world with data sciences*, Int Rob Auto J, 2(1) 00012, 2017

[3] Nannan Xi,Juho Hamari,Marc Riar : The challenges of entering the metaverse: An experiment on the effect of extended reality on workload,Article in Information Systems Frontiers,february 2022

[4] S.Cacciaguerra and G. D' Angelo, International Journal of Computer Games Technology,Volume 2008,Article ID 642314, Research Article - The Playing Session:Enhanced Playability for Mobile Gamers in Massive Metaverses

[5] Chris Collins("Fleep Tuque"),Looking to the Future:Higher Education in the Metaverse

[6] Qinglin Yang,Yetong Zhao,Huawei Huang,Zehui Xiong,Jiawen Kang - Fusing Blockchain and AI with Metaverse: A Survey, Guangdong University of Technology,China

[7] IEEE Transactions on Computational Social Systems,Vol 9 No.1,February 2022,MetaSocieties in Metaverse: Meta Economics and MetaManagement for MetaEnterprises and MetaCites

[8] GDA Capital ,Metaverse property, https://metaverse.properties/buy-in-decentraland/,December 21, 2021

[9] Zhao, X., Shi, C., You, X., and Zong, C. Analysis of mental workload : Are augmented and virtual reality consistent ? Frontiers in Psychology, 8, 71. https://doi.org/10. 3389/fpsyg.2017.0007

[10] Meta. Introducing meta: A social technology company,https://about.fb.com/news/2021/10/facebook-company-is-now-meta/

A Survey on Need of Green Computing

Iyrine George, Maria Benny and Sowmya P N
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala

Aparna S Balan
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501, Kerala
(email: aparna@vidyaacademy.ac.in)

***Abstract*—Green computing refers to the process of improving the efficiency of computing devices and reducing its negative impact on mankind and environment. Currently in networking field green computing technology has made a great impact in the environment, which makes it more efficient and proper use of network resources and less consumption of energy. Therefore, here it is discussed how green computing in cloud computing can overcome this problems and save the nature by its eco-friendly way of cloud computing. The term Green Computing refers to the practice of using energy wisely and efficiently, causing negligible degradation to environmental resources and parameters also maintaining feasibility on the other hand. In the present paper the authors have made a systematic study of several strategies, approaches and practices of green and energy efficient computing, in context to the growth and impact of the IT industry on environment in the recent years. So to make cloud computing more efficient in terms of energy consumption, the term Green Computing is used. Green Computing is basically responsible to use the computer and its related resources in an environment friendly way. In the present study the authors concentrate on the need for Green Technology and to make this globe sustainable for the coming years. At present green computing is under the consideration of businesses organizations and IT industries to improve environmental conditions for the better living of human being. This survey is attempt to discuss the need of green computing and steps required to be taken to improve the environmental condition in current era.**

***Index Terms*—Green computing, Cloud computing, Energy efficiency, Load balancing, Virtualization, Green IT**

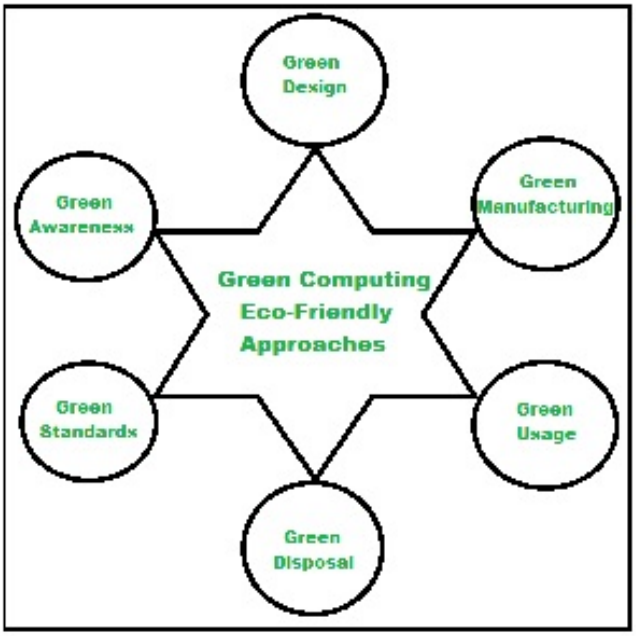


Fig. 1. Green Computing

I. Introduction

Green computing in Cloud computing is a way to reduce all the pollution caused by cloud computing. Green computing is the environmentally responsible and eco-friendly use of computers and their resources. It includes many benefits to the environment. The term Green Computing came into existence after the Energy Star Programme by the U.S. Environmental Protection Agency, in the year 1992. It was a voluntary labeling program, designed to promote energy-efficiency in hardware of all kinds. A similar program was also conducted in Europe and Asia later on. Cloud computing is basically defined as an "On Demand" service in which resources can be shared as per the need over the internet. In this present time, every business is heading itself to use cloud computing as it requires minimal management efforts. Due to this, the need to have more servers is increasing rapidly. More are the servers, more is the energy consumed by them. So to make cloud computing more efficient in terms of energy consumption, the term Green Computing is used. Cloud computing can be made more energy efficient by using technologies like workload consolidation and resource virtualization. Cloud computing is emerging as a new paradigm of large-scale distributed computing. Load balancing is one of the main challenges in cloud computing which is required to distribute the dynamic workload across multiple nodes to ensure that no single node is overwhelmed. The goal of load balancing is to minimize the resource consumption which will further reduce energy consumption and carbon emission rate that is the dire need of cloud computing. Cloud computing is the much sought after technology as it solves large computational problems by using various resources and also offers many other services through the internet. The computational offloading and the

increased demand for timely and appropriate response for real time applications has resulted in the need to improve the performance of the cloud. Currently in networking field green computing technology has made a great impact in the environment, which makes it more efficient and proper use of network resources and less consumption of energy. There has been tremendous growth in communication networks market both in infrastructure and wireless field. Green networking is the practice of selecting energy-efficient networking technologies and products, and minimizing resource use whenever possible. In the current trend, —going green has become an agenda for the IT industries in terms of public relations and reduced costs. The manufacturing industries contribute directly to pollution, whereas the IT industries have a hidden impact on environmental pollution caused by unconscious consumption of power and inefficient use of hardware devices. Hence, the main motto of green computing is to explore the reasons and find the solution of this indirect adverse effect on environment. Current trends of Green Computing are towards efficient utilization of resources. According to researchers in the past the focus was on computing efficiency and cost associated to IT equipments and infrastructure services were considered low cost and available. The future of Green Computing is going to be based on efficiency, rather than reduction in consumption. The primary focus on the Green IT is in the organization‘s self interest in energy cost reduction, Data Centers, and at Desktop. The secondary focus of Green IT needs to focus beyond energy use in the Data Center and the focus should be on innovation and improving alignment with overall corporate social responsibility efforts. This secondary focus will demand the development of Green Computing strategies.

II. Need of green computing to improve environmental condition

At present green computing is under the consideration of businesses organizations and IT industries to improve environmental conditions for the better living of human being.By green computing or green IT, we also mean research aimed at reducing the environmental impact of IT equipment: personal computers and servers, scanners, printers, smartphones and external peripherals. The term of Green Computing is the research and study that covers the computing device lifecycle from cradle to grave to conserve the energy.The main focus of Green technology is towards the improvement of manufacturing process, energy conservation and valuable recycling of computing devices. Computer systems are becoming progressively more omnipresent and a part of the global infrastructure, resulting in large installations of computing devices to provide multiple services. Today almost all streams like information technology, medical, education, transportation and agriculture uses computers and electronic devices which directly or indirectly causes the harm to environment. When we are using a computing device certainly some amount carbon dioxide is emitted which have obviously have a harmful effect on the users. The use of computers and others electronic devices increase Carbon footprint . The overall harms due to all these factors can be broadly divided into three categories that are:

- During manufacturing process.
- During uses of computing devices.
- During destruction process.

A. During manufacturing process:

During the Manufacturing process of computing devices different toxic materials like Lead, Mercury and Cadmium.Lead accumulates in the soil, having a high organic content is transferred to the top layers of the soil surface, where it may be retained for many years and cause harmful effects.Mercury is used in batteries, switches and printed circuit boards etc. Mercury is used in medical equipment, data transmission equipment, telecommunications equipment and cell phones as well.Apart from these three there exists a lot of toxic materials like hexavalent chromium, PVC plastic and brominated flame retardants, sulphur, carbon, silicon which are very harmful for human body and environment..

B. During uses of computing devices:

If we look at the energy consumption of the machines during the actual use and standby mode it is observed that during the standby mode the energy consumption ratio is more for example an inkjet printer uses twelve watts of energy during printing and take five watts during idle state.

C. During destruction process:

The scrap components of computers such as monitor, printers etc may contain toxic material such as lead, cadmium, beryllium etc. During the process of destruction and recycling of the computing devices different types of gases and liquids are produced from different toxic materials.The recycling and disposal of e-waste may involve significant risk to workers and communities because of unsafe exposure in recycling operations and leaking of materials such as heavy metals from landfills and incinerator ashes.

III. Energy Efficiency in Green Computing

Cloud computing can be made more energy efficient by using technologies like workload consolidation and resource virtualization. Cloud datacenter can reduce the energy consumed by server consolidation whereby different workloads can share same physical host with the concept of virtualization and the unused servers can be switched off in order to save energy. We will do the evaluation of green cloud on the basis of the linear power model. In the Linear power models we will use four different scheduling algorithms the Green scheduler, Green scheduler using virtual machines, Round-Robin scheduling using host and Round Robin using virtual machines. From the simulator we can select which algorithm we have to use at a particular instance. The number of servers and switches can be customized and the numbers of users are fixed.

The graphical comparison of the total energy consumed and server energy consumed is shown below.

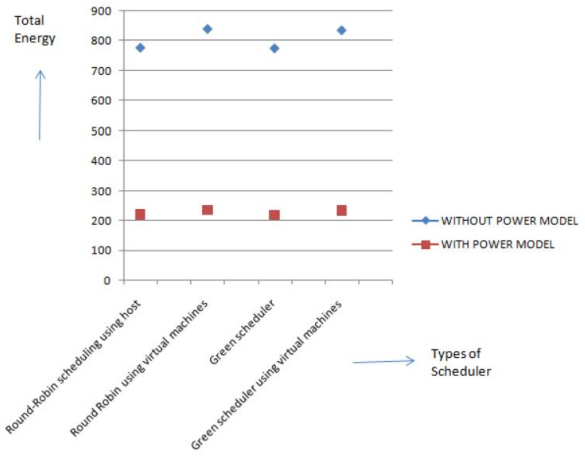


Fig. 2. Comparison for Total energy

IV. Green Computing in Cloud Computing

Green Cloud Computing is a model that gives positive side not only to nature but also to the data center and service provider as they utilize the resources efficiently. The positive side of this model is that it is eco-friendly, it is energy efficient, consolidation, virtualization etc.

1) **Eco-Friendly**
   Green Cloud computing is an eco-friendly model as it prevents cloud computing from harming the environment. It prevents the load on processor and reuse the resources that gives a benefit to use resources efficiently. The heat and all other harmful gases that used to get emerged from cloud computing is minimized due to this model.
2) **Energy Efficient**
   Green Computing in Cloud Computing also benefits us as well as nature by saving energy. It not let the power get wasted and use every energy to its max. Previously, the processors used to get hot and it doesn't work properly and energy was also wasted but green cloud computing solves this problem by saving energy and using it efficiently.
3) **Virtualization**
   Using some abstract process, this concept of virtualization lets run several logical computers in single physical computer. There are several software that let it done and process, like Hypervision, It coordinates with the underlying hardware components according to the instructions given by virtual machine.
4) **Consolidation**
   The process of deploying different data centers related data processing on a single server with virtual technology is consolidation. It implements the process level load balancing. It reduces the power consumption and saves energy.

V. Cloud Load Balancing Techniques

Cloud computing is emerging as a new paradigm of large-scale distributed computing. Load balancing is one of the main challenges in cloud computing which is required to distribute the dynamic workload across multiple nodes to ensure that no single node is overwhelmed. It helps in optimal utilization of resources and hence in enhancing the performance of the system. The goal of load balancing is to minimize the resource consumption which will further reduce energy consumption and carbon emission rate that is the dire need of cloud computing. Load balancing in clouds is a mechanism that distributes the excess dynamic local workload evenly across all the nodes. It is used to achieve a high user satisfaction and resource utilization ratio, making sure that no single node is overwhelmed, hence improving the overall performance of the system. Proper load balancing can help in utilizing the available resources optimally, thereby minimizing the resource consumption. It also helps in implementing fail-over, enabling scalability, avoiding bottlenecks and over-provisioning, reducing response time etc. Load balancing is also required to achieve Green computing in clouds which can be done with the help of the two factors: Reducing Energy Consumption and Reducing Carbon Emission.

VI. Green Computing in Communication Networks

As the new generation networking environment is taking shape, energy efficiency has pervaded the network infrastructure as a whole, to such extent as to become part of the network design criteria, and to carry across multiple networking domains for the achievement of a general target. A common opinion among green networking community is that the sole introduction of low consumption silicon technologies may not be enough to for drawing ahead current network equipment towards a greener networking environment. As a consequence, the problem of energy efficiency of networking environment is a concern as in other green computing areas. In the Infrastructural network, consumption of power, bandwidth and memory and resource utilization is the key issue. In large data center which is a facility used to house computer systems and associated components, such as telecommunications and storage systems the power consumption of these housed components are considered to start the energy saving process because everybody knows that data centers are becoming one of the major consumer of electricity in the industrialized world and the trend is set for it to rise even higher.

- **Power utilization in datacenters**
  One type of Green Computing focuses on reducing power consumption of computers. Most data centers, by design, consume vast amounts of energy in an incongruously wasteful manner, interviews and documents show. Online companies typically run their facilities at maximum capacity around the clock, whatever the demand. If we compare the average data center with the average office we can see that a data canter consumes around fifty four times more energy on a square meter by square meter basis.

- **Bandwidth shared among Datacenters**
  Data centre is the foremost ingredient of the cloud computing. NetShare, a mechanism for data center networks that requires no hardware changes to routers but allows bandwidth to be allocated predictably across services based on weights.
  The different bandwidth mechanisms are listed below:
  1) Fair-Share Bandwidth allocation
  2) TCP
  3) Bandwidth Capping
  4) Secondnet
  5) NetShare
  6) Approximate Fairness- Quantized Congestion
  7) Notification (AF-QCN)
- **Memory storage in cloud**
  Memory storage is one of main concern in cloud computing along with the other services like SaaS (Software as a Service), IaaS (Infrastructure as a Service), PaaS (Platform as a Service) and RaaS (Routing as a Service). The cloud keeps fixed free RAM size in virtual machine and increases or decreases RAM size when needed. It is observed that cloud service based memory utilization has been an effective technique for allocating the memories in a cloud computing environment.

VII. Impact of Green Computing in IT Industry

In the current trend, —going green has become an agenda for the IT industries in terms of public relations and reduced costs. Green computing focuses on the triple bottom line of economic viability, social responsibility, and environmental impact. It differs from traditional business practices that focus mainly on economic viability of a computing solution. IT industry has begun to address energy consumption in the data center through a variety of approaches including the use of more efficient cooling systems, virtualization, blade servers and storage area networks. But a basic problem persists in the path of achieving sustainability.

Green computing is a mindset that asks how we can satisfy the growing demand for network computing without putting such pressure on the environment. There is an alternative way to design a processor and a system such that we don't increase demands on the environment, but still provide an increased amount of processing capability to customers to satisfy their business needs. Green computing is not about going out and designing biodegradable packaging for products. Now the time has come to think about the efficiently use of computers and the resources which are non-renewable. It opens a new window for the new entrepreneur for harvesting with E-waste material and scrap computers. The greenest computer will not miraculously fall from the sky one day; it'll be the product of years of improvements.

VIII. Green Solutions: A Pilot Study on Green Technology and Green Computing

Green Computing primarily concerned about efficient use of computer power management, minimum discharge of carbon dioxide and to recycle e-waste to make the environment sustainable. Green Technology is one step above Green Computing where the people much more concerned about how to minimize any kind of technology related hazards and to make this globe safe for the use of coming generations. Due to the alarming degradation of the overall environmental situations over the whole world, it is very much needed that the researchers should propose some environmental technical solutions in more or less all fields of technology. Global warming is a serious issue today ,probably the most worrying threat to our planet at the present time is global warming. Green computing and the use of energy efficient green technologies can reduce this scaring impact of global warming. All we know the continuous use of non-recyclable energy resources can lead us to the dead end of the civilization.

IX. Computing Strategies for Green Cloud

Some of the Algorithmic Strategies are:

1) **First Come First Serve:**
   The task that arrives first will be scheduled first and the resources are allocated to it as per its need. Once the task is completed the next process in the queue is taken for execution.

2) **Round-Robin:**
   The task will be executed till the time called the time slice or the quantum that is set is reached. The task will then be put in the queue at the last and the task will again be taken when it reaches the front end and the remaining execution will happen.

3) **Min-Min Algorithm:**
   This algorithm is based on the minimum completion time and the small tasks are executed first and this incurs delay in the larger tasks or it may not schedule larger tasks.

4) **Max-Min Algorithm:**
   The tasks are analyzed and a matrix is formed representing the execution time. The tasks which has the maximum completion time is executed first by assigning the available resources to it.

5) **Swarm Optimization:**
   This provides intelligent multi-dimensional solutions forenergy-aware virtual machine consolidation. These are adaptive strategies where the load from over-used machines will be moved to machines that are lightly used. Generally, this provides a green solution by reducing the operational cost by effective task assignment. Swarm optimisation demonstrates to be more robust and highly efficient. Ant colony optimisation, Artificial bee colony, Particle swarm optimisation, Differential evolution are few methods which comes under this.

X. Current Trends

Current trends of Green Computing are towards efficient utilization of resources. Energy is considered as the main resource and the carbon footprints are considered the major threads to environment. Therefore the emphasis is to reduce the energy utilization and carbon footprints and increase the performance of Computing.

- **Energy Consumption:**
  Organizations are realizing that the source and amount of their energy consumption significantly contributes to Greenhouse Gas (GhG) emissions. In response to this finding, organizations are currently using the following equation: Reduced energy consumption = Reduced greenhouse gas emissions = Reduced operational costs for the data center.

- **E-Waste Recycling:**
  E-Waste is a manageable piece of the waste stream and recycling e-Waste is easy to adopt. Recycling computing equipment such as lead and mercury enables to replace equipment that otherwise would have been manufactured. The reuse of such equipments allows saving energy and reducing impact on environment, which can be due to electronic wastes.

- **Data Center:**
  Consolidation and Optimization Currently much of the emphasis of Green Computing area is on Data Centers, as the Data Centers are known for their energy hunger and wasteful energy consumptions.

- **Virtualization:**
  One of the main trends of Green Computing is virtualization. Abstraction of computer resources, such as the running two or more logical computer systems on one set of physical hardware is called virtualization. Virtualization is a trend of Green computing it offers virtualization software as well as management software for virtualized environments of computer resources.

- **IT Products and eco-labeling:**
  Another approach to promote Green Computing and save environment is to introduce policies all around the World, so that, companies design products to receive the eco-label.

XI. Conclusion

By practicing green cloud computing, the resources and equipments that are used in cloud computing are used efficiently. Current challenges to achieve Green Computing are enormous and the impact is on computing performance. All these efforts are still in limited areas and currently efforts are mainly to reduce energy consumption and e-Waste but the future of Green Computing will be depending on efficiency and Green products. We feel there are two aspects of green computing, one is spreading awareness in society about the benefits of green computing and another one is moving towards technological advancement in the manufacturing and recycling process to minimize the effects of toxic martial. Green computing represents a responsible way to address the issue of global warming. By adopting green computing practices, business leaders can contribute positively to environmental stewardship and protect the environment while also reducing energy and paper costs. Green computing represents a responsible way to address the issue of global warming. Therefore, there is a need to develop an Energy-efficient load balancing technique that can improve the performance of cloud computing along with maximum resource utilization, in turn reducing energy consumption as well as carbon emission to an extent that will help achieve Green Computing.

References

[1] Bei Guan, Jingzheng Wu, Yongji Wang and Samee U. Khan (2014), 'CIVSched: A Communication-Aware Inter-VM Scheduling Technique for Decreased Network Latency between Co-Located VMs', IEEE Transactions On Cloud Computing,Vol.2,No.3,pp.320-332.

[2] Dong Jiankang, Wang Hongbo, Cheng Shiduan (2015), 'Energy-Performance Tradeoffs in IaaS Cloud with Virtual Machine Scheduling',ChinaCommunications, Vol. 12, No. 2, pp. 155-166.

[3] Robert R. Harmon, Haluk Demirkan, The Corporate Sustainability Dimensions of Service-Oriented Information Technology, Annual SRII Global Conference, March 29 2011 – April 2 2011, DOI: 10.1109/SRII.2011.116,

[4] Prof. Riyaz A. Sheikh, Dr. U.A. Lanjewar , Green Computing- Embrace a Secure Future, International Journal of Computer Applications (0975 – 8887), volume 10-N.4, November 2010

[5] V. Nae, R. Prodan, and T. Fahringer, "Cost-Efficient Hosting and Load Balancing of Massively Multiplayer Online Games", Proceedings of the 11th IEEE/ACM International Conference on Grid Computing (Grid), IEEE Computer Society, October 2010, pages 9-17.

[6] A. M. Nakai, E. Madeira, and L. E. Buzato, "Load Balancing for Internet Distributed Services Using Limited Redirection Rates", 5th IEEE Latin-American Symposium on Dependable Computing (LADC), 2011, pages 156-165.

[7] Yan Chen, Shunqing Zhang, and Shugong Xu, —Fundamental Trade-offs on Green Wireless Networksǁ IEEE Communications Magazine • June 2011

[8] Green Computing is SMART COMPUTING – A Survey, Ms Swati Aggarwal, Mrs. Monika Garg, Mr. Pramod Kumar, International Journal of Engineering, Vol. 2, Issue 2, Feb 2012.

[9] —Green Computing - a new Horizon of Energy Efficiency and Electronic waste minimizationǁ: a Global Perspective, Shalabh Agarwal and Asoke Nath : Proceedings of IEEE CSNT-2011 held at SMVDU(Jammu) 03-06 June 2011, Page 688-693.

[10] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[11] Dick M, Drangmeister J, Kern E, Naumann S (2013) Green software engineering with agile methods in green and sustainable software (GREENS). In: 2013 2nd International Workshop,pp 78–85

[12] Adams WM (2006) The future of sustainability. Re-thinking environment and development in the twenty-first century: technical report, IUCN

[13] Pushtikant Malviya and Shailendra Singh, "A Study about Green Computing", International Journal of Advanced Research in Computer Science and Software Engineering International Journal of Advanced Research in Computer Science and Software Engineering, Volume3, Issue 6, Pages 790-794, June 2013

[14] S C Rachana, Dr. H S Guru Prasad, "Emerging Security Issues and Challenges in Cloud Computing", International Journal of Engineering Science and Innovative Technology, Volume 3, Issue 2,Pages 485-490, March 2014

Name Index

A

Adarsh K Sunny, 69
Adityalakshmi S Kumar, 95
Afra P L, 21
Aleena Vincent, 34
Amritha T A, 39
Anjali John, 1
Anjana K U, 1
Anjitha M P, 34
Anjitha N A, 128
Annmary M B, 107
Anupama P N, 27
Aparna S Balan 39,87,134
Arjun K, 64
Athira K M, 100
Athira P V, 81
Athul P S, 128
Aysha P A, 59

B

Blessy Joseph, 122

C

Catherine Joshi,128

D

Devika P H, 87
Dhilsha K P, 81
Dijesh P , 9,59,100
Dijina P D, 34
Dilna K Daison, 107
Diyana Ummer, 39
Dona James, 59
Dr Reji C Joy, 1,50,95

F

Frederic Sojan, 69

H

Hansa A M, 21
Haritha H, 100
Hridhya K B, 15

I

Iyrine George, 134

J

Jamseena N A, 95
Janish P S, 122
Jesly Johnson A, 1
Jineesha Joseph, 9
Jini James, 9
Josiya Mary, 39

K

K R Deepa, 50
Krishnapriya T S, 59

L

Lishel Raphael, 64

M

M R Aswathy, 122
M Shahil Haris, 116
Manasa K Mani, 100
Manesh D, 21,69,116
Manju P A, 9
Maria Benny, 134
Mariya C Johny, 75
Meghana K B, 87
Merin V Wilson, 64
Mohidh K M, 107
Mrudula M M, 95

N

Naina Vincent, 21

P

Praseetha Praveen, 116

R

Reshma Reghu, 27
Rohini T B, 69

S

Sajay K R, 15,64,107
Salkala K S, 34,81,128
Sandra Thomas, 81
Saranya V Simon, 75
Shejila K, 87
Siji K B, 27,75,122
Sibin Babu,50
Silpa Sasi, 75
Sowmya P N, 134
Sreeharsha V P, 50
Sruthy Krishna P U, 15
Steny Baby, 27
Suraj Kumar M, 116
Swetha M S, 15

VMCADS - 2022